LDS 09/457252

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 June 2002 (27.06.2002)

PCT

(10) International Publication Number
WO 02/50106 A2

(51) International Patent Classification[7]: C07K 14/195

(21) International Application Number: PCT/CA01/01847

(22) International Filing Date:
19 December 2001 (19.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/256,489 20 December 2000 (20.12.2000) US

(71) Applicant *(for all designated States except US)*: **PHAGETECH INC.** [CA/CA]; C.P. 387, Place du Parc, Montréal, Québec H2W 2N9 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PELLETIER, Jerry** [CA/CA]; 8 Lakeview, Baie D'Urfe, Quebec H9X 3B1 (CA). **GROS, Philippe** [CA/CA]; 107 Montrose, St. Lambert, Quebec J4R 1X4 (CA). **DUBOW, Michael** [CA/CA]; 4901 Coolbrook, Montreal, Quebec H3X 2K8 (CA).

(74) Agents: **DUBUC, J.** et al.; Goudreau Gage Dubuc, Stock Exchange Tower, Suite 3400, 800 Place Victoria, P.O. Box 242, Montréal, Québec H4Z 1E9 (CA).

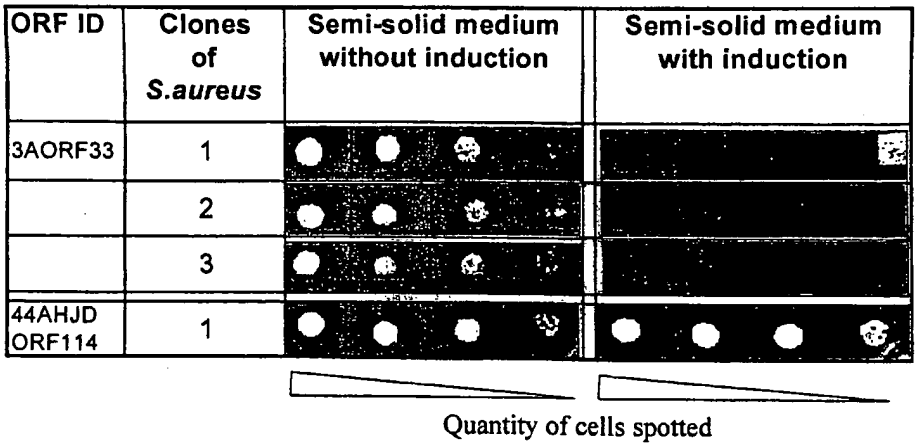
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS AUREUS GENE AND ITS ENCODED PROTEIN STAAU_R4



(57) Abstract: The invention relates to bacterial genes and proteins that are implicated in the process of fatty acid/phospholipid biosynthesis and also to bacteriophage genes and their protein products that interact with bacterial proteins implicated in fatty acid/phospholipid biosynthesis. More particularly, the invention relates to compositions and methods involving essential *Staphylococcus aureus* genes and its encoded proteins STAAU_R4. In addition, the invention relates to screening assays to identify compounds which modulate the level and/or activity of STAAU_R4 and to such compounds.

WO 02/50106 A2

TITLE OF THE INVENTION

COMPOSITIONS AND METHODS INVOLVING AN ESSENTIAL STAPHYLOCOCCUS AUREUS GENE AND ITS ENCODED PROTEIN STAAU_R4

FIELD OF THE INVENTION

The invention relates to bacterial genes and proteins that are implicated in the process of fatty acid/phospholipid biosynthesis and also to bacteriophage genes and their protein products that interact with bacterial proteins implicated in fatty acid/phospholipid biosynthesis. More particularly, the invention relates to compositions and methods involving essential *Staphylococcus aureus* genes and its encoded proteins STAAU_R4. In addition, the invention relates to screening assays to identify compounds which modulate the level and/or activity of STAAU_R4 and to such compounds.

BACKGROUND OF THE INVENTION

The Staphylococci make up a medically important genera of microbes known to cause several types of diseases in humans. *S. aureus* is a Gram positive organism which can be found on the skin of healthy human hosts and it is responsible for a large number of bacteremias.

S. aureus has been successfully treated with the penicillin derivative Methicillin in the past, but is now becoming increasingly resistant (MRSA - Methicillin Resistant *S. aureus*) to this antibiotic [Harbath *et al.*, (1998) Arch. Intern. Med. 158:182-189]. For example, *S. aureus* endocarditis mortality can range from 26-45%, and combined ß-lactam/aminoglycoside therapy is proving increasingly ineffective in disease eradication [Røder *et al.*, (1999) Arch. Intern. Med. 159:462-469].

It is no longer uncommon to isolate *S. aureus* strains which are resistant to most of the standard antibiotics, and thus there is an unmet medical need and demand for new anti-microbial agents, vaccines, drug screening

methods, and diagnostic tests for this organism. Antibiotics can be grouped into broad classes of activities against surprisingly few targets within the cell. Generally, the target molecule is a cellular protein that provides an essential function. The inhibition of activity of the essential protein leads either to death of the bacterial cell or to its inability to proliferate. Critical cellular functions against which antibiotics are currently in use include cell wall synthesis, folate and fatty acid metabolism, protein synthesis, and nucleic acid synthesis. Of particular relevance for the current invention is fatty acid/phospholipid biosynthesis.

A proven approach in the discovery of a new drug, referred to as target-based drug discovery to distinguish it from cell-based drug discovery, is to obtain a target protein and to develop *in vitro* assays to interfere with the biological function of the protein.

There remains a need to identify new bacterial targets and new target domains, and more particularly *S. aureus* bacterial targets which could be used to screen for and identify antibacterial and more particularly anti-*S. aureus* agents. There also remains a need to identify new antimicrobial agents, vaccines, drug screening methods and diagnosis and therapeutic methods.

The present invention seeks to meet these and other needs.

The present description refers to a number of documents, the content of which is herein incorporated by reference in their entirety.

SUMMARY OF THE INVENTION

The present invention relates to new antimicrobial agents, vaccines, drug screening methods and diagnosis and therapeutic methods.

More particularly, the invention relates to compounds which interact with STAAU_R4 and in particular to bacterial growth-inhibitory (or inhibitor) bacteriophage gene products that interacts with the *S. aureus* STAAU_R4 polypeptides.

The invention also relates to a pair of interaction proteins and parts or fragments thereof. More specifically, the invention relates to the

interacting domains of the *S. aureus* STAAU_R4 protein and to proteins which interact with same and block or inhibit a STAAU_R4 biological activity. In a particular embodiment, the invention relates to a pair of interacting domains comprised of that of STAAU_R4 and a polypeptide encoded by a bacteriophage ORF which specifically interacts therewith, such as the *S. aureus* bacteriophage 3A ORF 33. In one particularly preferred embodiment of the present invention, the interaction of these domains and a modulation thereof forms the basis for screening assays to identify modulators of STAAU_R4 biological function and more particularly of antimicrobials.

The present invention also relates to polynucleotides and polypeptides of a multiprotein complex believed to be involved in fatty acid/phospholipid biosynthesis replication containing STAAU_R4 as a subunit, as well as variants and portions thereof.

In another aspect, the invention relates to methods for using such polypeptides, polynucleotides, peptidomimetics and the like, including treatment and diagnosis of microbial diseases, amongst others.

In a further aspect, the invention relates to methods for identifying agonists and antagonists using the materials provided by the invention. In a related aspect, the invention relates to methods for treating microbial infections and conditions associated with such infections with the identified agonist or antagonist.

In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with microbial infections and conditions associated with such infections. In one embodiment, the diagnostic assay detects STAAU_R4 expression and/or activity.

In one particular embodiment of the invention, there is provided a method of identifying a compound that is active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2 or a biologically active fragment, or variant thereof, wherein SEQ ID NO: 2 or a biologically active fragment or variant thereof is capable of binding specifically with a polypeptide

comprising the sequence selected from SEQ ID NO: 4, a biologically active fragment thereof, and variant thereof, wherein the fragment or variant retains its capability of binding to SEQ ID NO:2, fragment, or variant thereof.

In one preferred embodiment of the invention, the identification of a compound active on a STAAU_R4 polypeptide is provided by a method comprising: contacting a first and a second polypeptide in the presence or absence of a candidate compound, wherein the first polypeptide comprises the amino acid sequence of SEQ ID NO: 2, a fragment or variant thereof which specifically binds to a second polypeptide derived from a bacteriophage ORF which is capable of binding specifically with one of SEQ ID NO: 2, a fragment, or variant thereof. In one particular embodiment, the second polypeptide is a phage ORF, a fragment thereof or variant thereof, wherein this second polypeptide maintains its biological activity; and detecting a biological activity of the first and/or second polypeptide, wherein a decrease in the biological activity in the presence thereof relative to the biological activity in the absence of the candidate compound identifies the candidate compound as a compound that is active on a polypeptide comprising the amino acid sequence of SEQ ID NO:2, fragment or variant thereof.

In one particular embodiment, the biological activity is the binding of the first and second polypeptides to each other, the method comprising: contacting an assay mixture comprising a) a first polypeptide which comprises the amino acid sequence of SEQ ID NO:2 or a biologically active fragment, or variant thereof, and b) a second polypeptide selected from the group consisting of SEQ ID NO: 4, a fragment thereof, and a variant thereof; with a test compound; measuring the binding of the first and the second polypeptides in the presence of the candidate compound relative to the binding in the absence thereof and; determining the ability of the candidate compound to interact with a STAAU_R4 polypeptide, fragment or variant thereof, wherein a decrease in the binding of the first and the second polypeptides in the presence of a candidate compound that interacts with the first polypeptide, relative to the binding in the absence thereof,

identifies the candidate compound as a compound that is active on a STAAU_R4 polypeptide, fragment or variant thereof. While in a preferred embodiment, the method identifies a compound active on STAAU_R4, the method can also identify compounds active on a bacteriophage ORF which specifically interacts with STAAU_R4.

In one embodiment, the step of detecting comprises the step of measuring the binding of the first and second proteins, wherein the first or the second protein is directly or indirectly detectably labeled.

The invention also encompasses a method of identifying an antimicrobial agent comprising determining whether a test compound is active on a *S. aureus* polypeptide, namely STAAU_R4 as set forth in SEQ ID NO: 2, or parts thereof.

In a further embodiment, identifying a compound active on a STAAU_R4 polypeptide is provided by a method which comprises: contacting a candidate compound with a polypeptide comprising the amino acid sequence of SEQ ID NO: 2; a fragment thereof, or a variant thereof, the fragment or variant retaining its biological activity, and detecting binding of the candidate compound thereto, wherein detection of binding is indicative that the compound is active on the polypeptide.

In different embodiments, the step of detecting includes measuring the binding of a candidate compound to the polypeptide, wherein the compound is directly or indirectly detectably labeled, by a method comprising, but not limited to, time-resolved fluorescence resonance energy transfer (TR-FRET), fluorescence polarization changes, measurement by surface plasmon resonance, scintillation proximity assay, biosensor assay, and phage display.

In one embodiment, a library of compounds is used. Non-limiting examples of candidate compound include a small molecule, a peptidomimetic compound, a peptide, and a fragment or derivative of a bacteriophage inhibitor protein.

In one embodiment, the candidate compound is a peptide synthesized by expression systems and purified, or artificially synthesized.

The invention further encompasses a method of identifying a compound that is active on a STAAU_R4 polypeptide, a fragment or a variant thereof, comprising the steps of contacting a candidate compound (or library thereof) with cells expressing a polypeptide comprising the amino acid sequence as set forth in SEQ ID NO: 2; and detecting STAAU_R4 activity in the cells, wherein a decrease in activity relative to STAAU_R4 activity in cells not contacted with a candidate compound is indicative of inhibition of STAAU_R4 activity. The invention also encompasses such a method but using a fragment or variant of SEQ ID NO:2.

Of course, the invention further encompasses methods of identifying a compound that modulates the activity of a STAAU_R4 polypeptide, wherein a compound increasing the activity relative to STAAU_R4 activity in cells not contacted with the candidate compound, is selected as a compound which is a stimulator of STAAU_R4 activity.

The invention further encompasses a method of identifying a compound that is active on a STAAU_R4 polypeptide, a fragment or a variant thereof, comprising the steps of contacting a candidate compound (or library thereof) in a cell-free assay, with a STAAU_R4 protein or biologically active portion thereof, either naturally occurring or recombinant in origin; and detecting STAAU_R4 activity, wherein a decrease in activity relative to STAAU_R4 activity in cell-free assay not contacted with a candidate compound is indicative of inhibition of STAAU_R4 activity.

The invention further encompasses an agonist or an antagonist of the activity of a STAAU_R4 polypeptide or a nucleic acid or gene encoding the polypeptide.

The assays described herein may be used as initial or primary screens to detect promising lead compounds for further development. The same assays may also be used in a secondary screening assay to measure the activity

of candidate compounds on a STAAU_R4 polypeptide. Often, lead compounds will be further assessed in additional, different screens. This invention also includes secondary STAAU_R4 screens which may involve biological assays utilizing *S. aureus* strains or other suitable bacteria.

Tertiary screens may involve the study of the effect of the agent in an animal. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, a test compound identified as described herein (e.g., a STAAU_R4 inhibiting agent, an antisense STAAU_R4 nucleic acid molecule, a STAAU_R4-specific antibody, or a STAAU_R4-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatment (e.g. bacterial infections), as described herein.

The invention further encompasses a method of making an antibacterial compound, comprising the steps of: a) determining whether a candidate compound is active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof, or a gene encoding the polypeptide; and b) synthesizing or purifying the candidate compound in an amount sufficient to provide a therapeutic effect when administered to an organism infected by a bacterium naturally producing a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof.

The invention further encompasses a method for inhibiting a bacterium, comprising contacting the bacterium with a compound active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof, or a nucleic acid encoding the polypeptide.

In one embodiment, the step of contacting is performed *in vitro*.

In another embodiment, the step of contacting is performed *in vivo* in an animal.

In another embodiment, bacterium is contacted with the active compound in combination with existing antimicrobial agents. Thus, the invention also relates to antimicrobial compositions comprising a compound of the present invention in combination with an existing antimicrobial agent. Of course, more than one active compound of the present invention could be combined with or without existing antimicrobial agent(s).

The invention further encompasses a method for treating or preventing a bacterial infection in an animal suffering from an infection or susceptible of suffering from same, comprising administering thereto a therapeutically or prophylactically effective amount of a compound active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, variant or fragment thereof, or nucleic acid sequence encoding same. The animal is preferably, but not necessarily a mammal, and more preferably a human. In one embodiment, the active compound is administred to the animal in combination with existing antimicrobial agents. Thus, the invention also relates to antimicrobial compositions comprising a compound of the present invention in combination with an existing antimicrobial agent.

The invention further encompasses a method of prophylactic treatment to prevent bacterial infection comprising contacting an indwelling device with a compound active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, variant or fragment thereof before its implantation into a mammal, such contacting being sufficient to prevent *S. aureus* infection at the site of implantation.

The invention further encompasses a method of prophylactic treatment to prevent infection of an animal by a bacterium comprising administering to the animal a prophylactically effective amount of a compound that is active on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, variant or fragment thereof or a gene encoding the polypeptide in an amount

sufficient to prevent infection of the animal. In a particular embodiment, the prophylactically effective amount reduces adhesion of the bacterium to a tissue surface of the mammal.

The invention further encompasses a method of diagnosing in an animal an infection with *S. aureus*, comprising: determining the presence in the animal of a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, part thereof, variant thereof, fragment thereof, epitope thereof or nucleic acid encoding same. Preferably the polypeptide is capable of specifically interacting with 3A ORF 33 . Preferably, the animal is a human.

In one embodiment, the determining step comprises contacting a biological sample of the animal or individual with an antibody specific for an epitope present on a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, variant or fragment thereof.

The invention further encompasses a method of diagnosing in an animal or individual an infection with *S. aureus*, comprising determining the presence in the animal or individual of a nucleic acid sequence encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, variant or fragment thereof, wherein the polypeptide is capable of specifically interacting with 3A ORF 33 .

In one embodiment, the determining step comprises contacting a nucleic acid sample of the animal or individual with an isolated, purified or enriched nucleic acid probe of at least 15 nucleotides in length that hybridizes under stringent hybridization conditions with the sequence of SEQ ID NO: 1, or the complement thereof.

The invention further encompasses an isolated, purified or enriched polynucleotide comprising a nucleotide sequence encoding a polypeptide, which can interact with a bacterial growth-inhibitory (or inhibitor) bacteriophage 3A ORF 33 gene product or part thereof.

In one particular embodiment, the isolated, purified or enriched polynucleotide comprises a nucleotide sequence encoding a polypeptide corresponding to SEQ ID NO: 2, or complement thereof.

In another particular embodiment of the present invention, the isolated, purified or enriched polynucleotide comprises a nucleotide sequence having at least 60 %, at least 70 %, at least 80 %, and more preferably at least 90% identity to the sequence of SEQ ID NO: 1, or to the complement thereof.

The invention further encompasses an isolated, purified or enriched polypeptide comprising the amino acid sequence of SEQ ID NO: 2, a variant or fragment thereof.

The invention further encompasses an isolated, purified or enriched polypeptide comprising or consisting in the amino acid sequence of SEQ ID NO: 2, or a fragment or variant thereof.

In one particular embodiment, the isolated, purified or enriched polypeptide comprises or consists of an amino acid sequence having at least 50%, at least 55%, at least 60%, at least 70%, and more preferably at least 80%, at least 90%, at least 95% or at least 99% identity to the amino acid sequence of SEQ ID NO: 2.

In one particular embodiment, the isolated, purified or enriched polypeptide of the present invention comprises or consists of an amino acid sequence having least 60%, at least 70%, at least 75%, at least 80%, more preferably at least 90%, and more preferably at least 95% or at least 99% similarity to the amino acid sequence of SEQ ID NO: 2.

In one particular embodiment, the sequence of SEQ ID NO:2 or fragment or variant thereof is part of a chimeric protein.

The invention further encompasses an isolated, purified or enriched antibody specific for an epitope encoded by the amino acid sequence set forth in SEQ ID NO: 2.

The invention further encompasses a composition comprising two polypeptides, a bacteriophage-encoded polypeptide and a *S. aureus*

STAAU_R4 polypeptide corresponding to SEQ ID NO: 2, or variant or fragment thereof. In another embodiment, the invention encompasses a composition comprising two interacting polypeptides derived from a bacteriophage encoded polypeptide and a *S. aureus* STAAU_R4 polypeptide. As such, the invention encompasses a composition comprising two nucleic acid sequences encoding these directly interacting polypeptides.

In another embodiment, the invention encompasses a composition comprising a pair of specifically interacting domains, the pair comprising: a STAAU_R4 polypeptide and a polypeptide encoded by a bacteriophage ORF which specifically interacts with the STAAU_R4 polypeptide.

Further, the invention encompasses a process for producing a pharmaceutical composition comprising: a) carrying out a screening assay of the present invention aimed at identifying a compound that is active on a STAAU_R4 polypeptide comprising the amino acid sequence of SEQ ID NO: 2, biologically active fragment thereof, or variant thereof; and b) mixing the compound identified in a) with a suitable pharmaceutical carrier.

In a particular embodiment, the process for producing a pharmaceutical composition comprises a) identifying a compound that is active on a STAAU_R4 polypeptide comprising the amino acid sequence of SEQ ID NO: 2, biologically active fragment thereof, or variant thereof that binds specifically with a second polypeptide derived from a bacteriophage ORF; and b) mixing the compound identified in a) with a suitable pharmaceutical carrier.

In a further embodiment of this process of producing a pharmaceutical composition, the process further includes a scaling-up of the preparation for isolating of the identified compound active on the STAAU_R4 polypeptide. In yet another embodiment of this process of producing a pharmaceutical composition, the pharmaceutical composition prepared comprises a derivative or homolog of the compound identified in a).

Also, the invention encompasses the use of one of: a) a STAAU_R4 polypeptide comprising the amino acid sequence of SEQ ID NO: 2,

a biologically active fragment thereof or variant thereof, b) a composition comprising a pair of specifically interacting domains comprised of a polypeptide of STAAU_R4, biologically active fragment thereof or variant thereof and a polypeptide encoded by a bacteriophage ORF which specifically interacts with STAAU_R4; or c) an assay mixture comprising a first polypeptide which comprises the amino acid sequence of SEQ ID NO:2, biologically active fragment thereof or variant thereof and a second polypeptide encoded by a bacteriophage ORF which specifically interact with each other; for the identification of a compound that is active on a polypeptide comprising the amino acid sequence of SEQ ID NO:2, biologically active fragment thereof or variant thereof.

In yet another embodiment of the present invention, there is provided an isolated polypeptide or polynucleotide sequence which comprises a bacteriophage ORF interacting domain which enables specific binding of the encoded bacteriophage ORF interacting domain to a bacteriophage ORF, the sequence being mutagenized in the non-bacteriophage ORF interacting domain.

Further features and advantages of the invention will become more fully apparent in the following description of the embodiments and drawings thereof, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which:

Fig. 1 shows the nucleotide (SEQ ID NO: 1) and amino acid (SEQ ID NO: 2) sequences of *S. aureus* STAAU_R4.

Fig. 2 shows the nucleotide (SEQ ID NO: 3) and the amino acid (SEQ ID NO: 4) sequences of *S. aureus* bacteriophage 3A ORF 33.

Fig. 3 shows the bacterial inhibitory potential of bacteriophage 3A ORF 33 and the expression vector used to induce its expression in *S. aureus*. A) Schematic diagram of expression vector pTM/ORF used to induce expression

of 3A ORF 33 in *S. aureus* cells; B) Results of a screen to assess the inhibitory potential of 3A ORF 33 when expressed in *S. aureus* grown on semi-solid support media; and C), D) Results of assays for inhibitory potential of 3A ORF 33 when expressed in *S. aureus* grown in liquid medium followed by plating on semi-solid medium either containing (C) or not containing (D) the antibiotic necessary to maintain the selective pressure for the plasmid.

Fig. 4 shows affinity chromatography using GST/3A ORF 33 ligands with a 10.0 mg/ml *S. aureus* extract. Eluates from affinity columns containing the ligands at 0, 0.1, 0.5, 1.0, and 2.0 mg/ml resin were resolved by 14% SDS-PAGE and the gel was stained with silver nitrate. Micro-columns were sequentially eluted with 100 mM ACB containing 0.1% Triton X-100 (SDS-PAGE not shown), 1 M NaCl ACB, and 1% SDS. Each molecular weight marker is approximately 200 ng. The lanes labeled ACB indicate eluates from a 2.0 mg/ml ligand column loaded only with ACB buffer containing 100 mM NaCl. The arrow designated PT40 indicates the position of the band that was excised for protein identification.

Fig. 5 shows affinity chromatography using 3A ORF 33 fused to a polyhistidine sequence (MSC2/3A ORF 33) as ligand with a 10.0 mg/ml *S. aureus* extract. Eluates from affinity columns containing the ligands at 0, 0.1, 0.5, 1.0, and 2.0 mg/ml resin were resolved by 14% SDS-PAGE and the gel was stained with silver nitrate. Micro-columns were sequentially eluted with 100 mM ACB containing 0.1% Triton X-100 (SDS-PAGE not shown), 1 M NaCl ACB, and 1% SDS. Each molecular weight marker is approximately 200 ng. The lanes labeled ACB indicate eluates from a 2.0 mg/ml ligand column loaded only with ACB buffer containing 100 mM NaCl. The arrow designated PT40 indicates the position of the band that was excised for protein identification.

Fig. 6 shows affinity chromatography using control GST as ligand with a 10.0 mg/ml *S. aureus* extract. Eluates from affinity columns containing the ligands at 0, 0.1, 0.5, 1.0, and 2.0 mg/ml resin were resolved by 14% SDS-PAGE and the gel was stained with silver nitrate. Micro-columns were

sequentially eluted with 100 mM ACB containing 0.1% Triton X-100 (SDS-PAGE not shown), 1 M NaCl ACB, and 1% SDS. Each molecular weight marker is approximately 200 ng. The lanes labeled ACB indicate eluates from a 2.0 mg/ml ligand column loaded only with ACB buffer containing 100 mM NaCl.

Fig. 7 shows the tryptic peptide mass spectrum analysis of the PT40 protein that interacted with 3A ORF 33 in affintity chromatography. The PT40 band was identified as an open reading frame, herein referred as STAAU_R4, found in Contig782 of the University of Oklahoma genome sequencing project database (http://www.genome.ou.edu/staph.html).

Fig. 8 shows the results of the BLAST searching analysis of publically available databases including NCBI (nr). STAAU_R4 is significantly similar to PlsX protein from a variety of bacteria, including *B. subtilis* (gi|6686325|sp|P71018|PLSX_BACSU: FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLSX).

Fig. 9 shows the global optimal alignment of the amino acid sequences of STAAU_R4 and *B. subtilis* PlsX and the identification of the start codon of the *S. aureus* gene. A) Result of the global optimal alignment reveals a 53% identity between STAAU_R4 and PlsX of *B. subtilis*; B) Results of the analysis of the *S. aureus* stop codon-to-stop codon DNA region containing STAAU_R4 revealing the presence of an ATG start codon associated with a predicted ribosomal binding site (RBS) sequence (in bold on the nucleotide sequence).

Fig. 10 shows A) an overview of the cloning of *S. aureus* STAAU_R4 and B) an overview of the cloning of phage 3A ORF 33 in the pGEX-6PK vector; C) the vector pGEX-6PK containing GST in fusion with PreScission protease cleavage site and the heart muscle kinase (HMK) phosphorylation site, followed by *Bam*HI,*Eco*RI and SalI cloning sites; and D) the results of the Far western affinity blotting analysis that was designed to test the interaction between purified 3A ORF 33 and [^{32}P]-ATP-radiolabelled STAAU_R4 polypeptides. Lanes 1 to 4 represent increasing amount (0.125, 0.250, 0.5 and 1.0 ug) of 3A ORF 33

resolved on SDS-polyacrylamide gel electrophoresis and blotted onto nitrocellulose membrane.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the accompanying drawing which is exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

The invention relates to the discovery of an essential gene and its encoded polypeptide in *S. aureus* and portions thereof useful for example in screening, diagnostics, and therapeutics. More specifically, the invention also relates to *S. aureus* STAAU_R4 polypeptides and polynucleotides as described in greater detail below, and to a pair of polynucleotides encoding a pair of interacting polypeptides, to the pair of polypeptides themselves, or interacting domains thereof. In a particular embodiment, the pair includes a *S. aureus* STAAU_R4 polypeptide or interacting domain thereof and a 3A ORF 33 or interacting domain thereof. In one embodiment, the invention relates to STAAU_R4 having the nucleotide or amino acid sequence disclosed as SEQ ID NO: 1 or SEQ ID NO: 2, respectively. The sequences presented as SEQ ID NOs: 1 and 2 represent an exemplification of the invention, since those of ordinary skill will recognize that such sequences can be usefully employed in polynucleotides in general, including ribopolynucleotides.

The methodology of two previous inventions (U.S. Provisional Patent Application 60/110,992, filed December 3, 1998, and PCT International Application WO1999/IB99/02040, filed December 3, 1999) has been used to identify and characterize essential polynucleotide and polypeptide sequences from *S. aureus*.

Thus, in a particular embodiment of the present invention, there is provided polynucleotide and polypeptide sequences isolated from *S.*

aureus that can be used in a drug screening assay to identify compounds with anti-microbial activity. The polynucleotide and polypeptide sequences can be isolated using a method similar to those described herein, or using another method. In addition, such polynucleotide and polypeptide sequences can be chemically synthesized. The identification of the *S. aureus* STAAU_R4 sequence as a target for a bacteriophage validates the approach of the present invention to identify bacterial targets and also validates STAAU_R4 as a key target for antibacterial drug development as well as diagnosis and treatment methods based thereon.

DEFINITIONS

In order to provide a clear and consistent understanding of terms used in the present description, a number of definitions are provided hereinbelow.

The terminology "active on", with reference to a particular cellular target, such as the product of a particular gene, means that the target is an important part of a cellular pathway which includes that target and that an agent or compound acts on that pathway. Thus, in some cases the agent or compound may act on a component upstream or downstream of the stated target (i.e. indirectly on the target), including a regulator of that pathway or a component of that pathway. In general, an antibacterial agent is active on an essential cellular function, often on a product of an essential gene (i.e. directly on the target).

The terminology "active on" also refers to a measurable effect of the compound on the target it is active on (as compared to the activity of the target in the absence of the compound). The activity referred thereto is any one of a biological activity of one of the polypeptides of the present invention.

As used herein, the terms "inhibit", "inhibition", 'inhibitory", and "inhibitor" all refer to a function of reducing a biological activity or function. Such reduction in activity or function can, for example, be in connection with a cellular component (e.g., an enzyme), or in connection with a cellular process (e.g.,

synthesis of a particular protein), or in connection with an overall process of a cell (e.g., cell growth). In reference to cell growth, the inhibitory effects may be bacteriocidal (killing of bacterial cells) or bacteriostatic (i.e. - stopping or at least slowing bacterial cell growth). The latter slows or prevents cell growth such that fewer cells of the strain are produced relative to uninhibited cells over a given time period. From a molecular standpoint, such inhibition may equate with a reduction in the level of, or elimination of, the transcription and/or translation and/or stability of a specific bacterial target(s), and/or reduction or elimination of activity of a particular target biomolecule.

As used herein, the terminologies "STAAU_R4 polypeptide", "plsX polypeptide" or the like refer to a polypeptide encompassing *S. aureus* STAAU_R4 (SEQ ID NO: 2), variant thereof or an active domain of *S. aureus* STAAU_R4. As used herein, the term "active domain of *S. aureus* STAAU_R4", "biologically active polypeptide of STAAU_R4" or the like refers to a polypeptide fragment or portion of *S. aureus* STAAU_R4 that retains an activity of *S. aureus* STAAU_R4. The term "STAAU_R4 polypeptide" is meant to encompass *S. aureus* STAAU_R4 or an active domain of *S. aureus* STAAU_R4 that is fused to another, non-STAAU_R4 polypeptide sequence.

"STAAU_R4 activity" "polypeptide comprising the amino acid sequence SEQ ID NO: 2 activity" "plsX polypeptide activity" or "biological activity" of STAAU_R4 or other polypeptides of the present invention is defined as a detectable biological activity of a gene, nucleic acid sequence, protein or polypeptide of the present invention. This includes any physiological function attributable to the specific biological activity of STAAU_R4, or phage ORF of the present invention. Non-limiting examples of the biological activities may be made directly or indirectly. STAAU_R4 biological activity, for example, is not limited, however, to its function in fatty acid/phospholipid biosynthesis. Biological activities may also include simple binding to other factors (polypeptides or otherwise), including compounds, substrates, and of course interacting proteins. Thus, for STAAU_R4, biological activity includes any standard biochemical measurement

of STAAU_R4 such as conformational changes, phosphorylation status or any other feature of the protein that can be measured with techniques known in the art. STAAU_R4 biological activity also includes activities related to STAAU_R4 gene transcription or translation, or any biological activities of such transcripts or translation products. The instant invention is also concerned with STAAU_R4 interaction with an inhibitory polypeptide of the present invention, biological activity of STAAU_R4 and fragment thereof also includes assays which monitor binding and other biochemical measurements of these polypeptides. Furthermore, for certainty, the terminology "biological activity" also includes measurements based on the interaction of domains of interacting proteins of the present invention (e.g. the phage ORF interacting domain of STAAU_R4, or STAAU_R4 interacting domain of a phage ORF domains thereof). Non-limiting examples of "biological activity" include one or more of the following:

i) Binding to a bacterial growth inhibitory ORF derived from a bacteriophage including a 3A ORF 33 polypeptide or part thereof.

Determining the binding between polypeptides of the present invention can be accomplished by one of the methods described below or known in the art for determining direct binding. While it might be advantageous in certain embodiments of the present invention to provide a binding assay which is amenable to automation and more particularly to high-throughput, the present invention is not so limited. The binding or physical interaction between a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, provided herein, or fragment thereof and a bacteriophage protein 3A ORF 33 or portion thereof may be between isolated polypeptides consisting essentially of the sequence necessary for binding, or, alternatively, the respective polypeptide sequence may be comprised within a larger polypeptide.

A number of non-limiting methods, useful in the invention, to measure the binding of bacteriophage 3A ORF 33 to a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or fragment thereof are described below. Binding can be measured by coupling one molecule to a surface or support such

as a membrane, a microtiter plate well, or a microarray chip, and monitoring binding of a second molecule by any number of means including but not limited to optical spectroscopy, fluorometry, and radioactive label detection.

For example, Time-Resolved Fluorescence Resonance Energy Transfer (TR-FRET), in which the close proximity of two fluorophores, whether intrinsic to, as in the case of a naturally-fluorescent amino acid residue such as tryptophan, or either covalently or non-covalently bound to a separate molecule, causes the emission spectrum of one fluorophore to overlap with the excitation spectrum of the second, and thus dual fluorescence following excitation of only one fluorophore is indicative of binding. An additional assay useful in the present invention is fluorescence polarization, in which the quantifiable polarization value for a given fluorescently-tagged molecule is altered upon binding to a second molecule. Surface plasmon resonance assays can be used as a quantitative method to measure binding between two molecules by the change in mass near an immobilized sensor caused by the binding of one protein from the aqueous phase to a second immobilized on the sensor. A scintillation proximity assay can also be used to measure binding of a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, and fragment thereof and a bacteriophage ORF or fragment thereof in which binding in the proximity to a scintillant converts radioactive particles into a photon signal that is detected by a scintillation counter or other detector. Additionally, binding can be evaluated by a Bio Sensor assay, which is based on the ability of the sensor to register changes in admittance induced by ion-channel modulation following binding. Phage display is also a powerful quantitative assay to measure protein:protein interaction using colourimetric ELISA (enzyme-linked immunosorbent assay).

ii) The complementation of glycerol-3-phosphate auxotrophy

A method, useful in the invention, is to measure the ability of a STAAU_R4 polypeptide to relieve glycerol-3-phosphate auxotrophy of an *E. coli* strain that possesses the *plsB26* and *plsX50* mutations. A strain of *E. coli* bearing mutations in two distinct loci, *plsB* and *plsX*, depends upon the presence of

glycerol-3-phosphate in the medium for its growth. Introduction, into this strain, of a polynucleotide sequence encoding wild-type PlsX circumvents the glycerol-3-phosphate growth requirement. Strains are assessed for the relief of their auxotrophic requirement by plating, in parallel, onto minimal medium containing or not containing glycerol-3-phosphate. Growth on the plates lacking glycerol-3-phosphate indicates complementation of the auxotrophic phenotype.

This assay has proven useful to assess whether homologues of *E. coli plsX* from other bacterial species, including the Gram-negative bacterium *Salmonella typhimurium* and the Gram-positive bacteria *Clostridium butyricum* and *Bacillus subtilis* can functionally replace PlsX in *E. coli*. Similar assays can be developed to test the activity of PlsX from *S. aureus* and to screen for inhibitors of such activity.

As used herein, the term "polynucleotide encoding a polypeptide" or equivalent language encompasses polynucleotides that include a sequence encoding a polypeptide of the invention, particularly a bacterial polypeptide and more particularly a polypeptide of *S. aureus* STAAU_R4 protein having an amino acid sequence set out in Fig. 1, SEQ ID NO: 2. The term also encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, polynucleotides interrupted by integrated phage, an integrated insertion sequence, an integrated vector sequence, an integrated transposon sequence, or otherwise altered due to RNA editing or genomic DNA reorganization) together with additional regions that also may contain coding and/or non-coding sequences.

As used herein, the term "STAAU_R4 gene" "PlsX gene" is meant to encompass a polynucleotide encoding a *S. aureus* STAAU_R4 polypeptide. Any additional nucleotide sequences necessary to direct transcription of RNA encoding a *S. aureus* STAAU_R4 polypeptide, either in a cell or *in vitro*, will be termed "regulatory sequences", which include but are not limited to transcriptional promoters and enhancers, and transcription terminators.

As used herein, the term "ORF 33" or "phage 3A ORF 33" or "3A ORF 33" encompasses a polynucleotide comprising or consisting of the sequence provided in Fig. 2 (SEQ ID NO: 3), which encodes a gene product known as the 3A ORF 33 gene product.

As used herein, the term "polynucleotide(s)" generally refers to any polyribonucleotide or poly-deoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotide(s)" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-, double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded regions, or a mixture of single- and double-stranded regions. In addition, "polynucleotide" as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term "polynucleotide(s)"also includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotide(s)" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term "polynucleotide(s)" as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including, for example, simple and complex

cells. "Polynucleotide(s)" also embraces short polynucleotides often referred to as oligonucleotide(s). Polynucleotides can also be DNA and RNA chimeras.

As used herein, the term "polypeptide(s)" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds. "Polypeptide(s)" refers to both short chains, commonly referred to as peptides, oligopeptides and oligomers and to longer chains generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptide(s)" include those modified either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. Modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation, selenoylation, sulfation and transfer-RNA mediated addition of amino acids to proteins, such as arginylation, and ubiquitination. See, for instance: Proteins – Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); Wold, F., Posttranslational

Protein Modifications: Perspectives and Prospects, pgs. 1-12 in Posttranslational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., Meth. Enzymol. 182:626-646 (1990); and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62(1992). Polypeptides may be branched or cyclic, with or without branching. Cyclic, branched and branched circular polypeptides may result from post-translational natural processes and may be made by entirely synthetic methods, as well.

As used herein, the term "variant(s)" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, respectively, but retains one or more of the biological activities of the initial (e.g. non-variant) polynucleotide or polypeptide of the present invention (e.g. STAAU_R4). A typical variant of a polynucleotide differs in nucleotide sequence from another reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, and truncations in the polypeptide encoded by the reference sequence, or in the formation of fusion proteins, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. The present invention also includes variants of each of the polypeptides of the invention, that is polypeptides that vary from the referents by conservative amino acid substitutions whereby a residue is substituted by another with like characteristics. Typically, such substitutions are among Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; and among the basic residues Lys and Arg; or aromatic residues Phe and

Tyr. Particularly preferred are variants in which 1-10, 1-5, 1-3, 2-3, or 1 amino acid or amino acids are substituted, deleted, or added in any combination. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques, by direct synthesis, and by other recombinant methods known to skilled artisans. As used herein, the term "fragment", when used in reference to a polypeptide, is a polypeptide having an amino acid sequence that is entirely the same as part but not all of the amino acid sequence of the polypeptide according to the invention from which it "derives". As with *S. aureus* STAAU_R4 polypeptides, fragments may be "free-standing" ("consisting of"), or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region in a single larger polypeptide.

The term "isolated", when used in reference to a nucleic acid means that a naturally occurring sequence has been removed from its normal cellular (*e.g.*, chromosomal) environment or is synthesized in a non-natural environment (*e.g.*, artificially synthesized). Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90-95% pure at least) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

The term "enriched", when used in reference to a polynucleotide means that the specific DNA or RNA sequence constitutes a significantly higher fraction (2-5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in cells from which the sequence was originally taken. This could be caused by a person, by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that enriched does not imply

that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased.

As used herein, the term "significantly higher fraction" indicates that the level of enrichment is useful to the person making such an enrichment and indicates an increase in enrichment relative to other nucleic acids of at least about 2-fold, or 5- to 10-fold or even more. The term also does not imply that there is no DNA or RNA from other sources. The other source of DNA may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occurring events, such as viral infection, or tumor type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

As used herein, the term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation). Instead, it represents an indication that the sequence is relatively more pure than in the natural environment (compared to the natural level, this level should be at least 2-5 fold greater, *e.g.*, in terms of mg/mL). Individual clones isolated from a genomic or cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 10^6-fold purification of the native message over its proportion in naturally occurring cells. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of

magnitude is expressly contemplated. A genomic library can be used in the same way and yields the same approximate levels of purification.

The terms "isolated", "enriched", and "purified" used with respect to nucleic acids, above, may similarly be used to denote the relative purity and abundance of polypeptides. These, too, may be stored in, grown in, screened in, and selected from libraries using biochemical techniques familiar in the art. Such polypeptides may be natural, synthetic or chimeric and may be extracted using any of a variety of methods, such as antibody immunoprecipitation, other "tagging" techniques, conventional chromatography and/or electrophoretic methods. Some of the above utilize the corresponding nucleic acid sequence.

As used herein, the term "complement" when used in reference to a given polynucleotide sequence refers to a sequence of nucleotides which can form a double-stranded heteroduplex in which every nucleotide in the sequence of nucleotides is base-paired by hydrogen bonding to a nucleotide opposite it in the heteroduplex with the given polynucleotide sequence. The term may refer to a DNA or an RNA sequence that is the complement of another RNA or DNA sequence. As used herein, the term "hybridizes" refers to the formation of a hydrogen-bonded heteroduplex between two nucleic acid molecules. Generally, a given nucleic acid molecule will hybridize with its complement, or with a molecule that is sufficiently complementary to the given molecule to permit formation of a hydrogen-bonded heteroduplex between the two molecules.

As used herein, the term "probe" refers to a polynucleotide of at least 15 nucleotides (nt), 20 nt, 30 nt, 40 nt, 50 nt, 75 nt, 100 nt, 200 nt, 500 nt, 1000 nt, and even up to 5000 to 10,000 nt in length.

"Identity" and "similarity," as used herein and as known in the art, are relationships between two or more polypeptide sequences or two or more polynucleotide sequences, as the case may be, as determined by comparing the sequences.

Amino acid or nucleotide sequence "identity" and "similarity" are determined from an optimal global alignment between the two sequences

being compared. A non-limiting example of optimal global alignment can be carried-out using the Needleman - Wunsch algorithm [Needleman and Wunsch, 1970, J. Mol. Biol. 48:443-453]. "Identity" means that an amino acid or nucleotide at a particular position in a first polypeptide or polynucleotide is identical to a corresponding amino acid or nucleotide in a second polypeptide or polynucleotide that is in an optimal global alignment with the first polypeptide or polynucleotide. In contrast to identity, "similarity" encompasses amino acids that are conservative substitutions.

The term "conservative" substitution is well-known in the art and broadly refers to a substitution which does not significantly change the chemico-physical properties of the substituted amino acid. For example, a "conservative" substitution is any substitution that has a positive score in the blosum62 substitution matrix [Hentikoff and Hentikoff, 1992, Proc. Natl. Acad. Sci. USA 89: 10915-10919]. By the statement "sequence A is n% similar to sequence B" is meant that n% of the positions of an optimal global alignment between sequences A and B consists of conservative substitutions. By the statement "sequence A is n% identical to sequence B" is meant that n% of the positions of an optimal global alignment between sequences A and B consists of identical residues or nucleotides. Optimal global alignments in this disclosure used the following parameters in the Needleman-Wunsch alignment algorithm:

For polypeptides:

Substitution matrix: blosum62.

Gap scoring function: -A -B*LG, where A=11 (the gap penalty), B=1 (the gap length penalty) and LG is the length of the gap.

For nucleotide sequences:

Substitution matrix: 10 for matches, 0 for mismatches.

Gap scoring function: -A -B*LG where A=50 (the gap penalty), B=3 (the gap length penalty) and LG is the length of the gap.

The term ' identity' and 'similarity' between sequences can be extended to their fragments. An optimal local alignment between sequences A

and B is the highest scoring alignment of fragments of A and B. By the statement "sequence A is n% similar locally to B" is meant that n% of the positions of an optimal local alignment between sequences A and B consists of conservative substitutions. By the statement "sequence A is n% identical locally to B" is meant that n% of the position of an optimal local alignment between sequences A and B consists of identical residues or nucleotides. A non-limiting example of optimal local alignment can be carried-out using the Smith-Waterman algorithm [Smith, T.F and Waterman, M.S. 1981. Identification of common molecular subsequences. J. Mol. Biol. 147:195-197].

Of course, the above-listed parameters are but one specific example of alignment algorithm parameters. Numerous algorithms and parameters are available and known to the person of ordinary skill.

Typical conservative substitutions are among Met, Val, Leu and Ile; among Ser and Thr; among the residues Asp, Glu and Asn; among the residues Gln, Lys and Arg; or aromatic residues Phe and Tyr. In calculating the degree (most often as a percentage) of similarity between two polypeptide sequences, one considers the number of positions at which identity or similarity is observed between corresponding amino acid residues in the two polypeptide sequences in relation to the entire lengths of the two molecules being compared.

As used herein, the term "antibody" is meant to encompass constructions using the binding (variable) region of such an antibody, and other antibody modifications. Thus, an antibody useful in the invention may comprise a whole antibody, an antibody fragment, a polyfunctional antibody aggregate, or in general a substance comprising one or more specific binding sites from an antibody. The antibody fragment may be a fragment such as an Fv, Fab or $F(ab')_2$ fragment or a derivative thereof, such as a single chain Fv fragment. The antibody or antibody fragment may be non-recombinant, recombinant or humanized. The antibody may be of an immunoglobulin isotype, e.g., IgG, IgM, and so forth. In addition, an aggregate, polymer, derivative and conjugate of an immunoglobulin or a fragment thereof can be used where appropriate. Neutralizing antibodies are

especially useful according to the invention for diagnostics, therapeutics and methods of drug screening and drug design.

As used herein, the term "specific for an epitope present on a *S. aureus* STAAU_R4 polypeptide", when used in reference to an antibody, means that the antibody recognizes and binds an antigenic determinant present on a *S. aureus* STAAU_R4 polypeptide or fragment thereof according to the invention.

As used herein, the term "antigenically equivalent derivative(s)" encompasses a polypeptide, polynucleotide, or the equivalent of either which will be specifically recognized by certain antibodies which, when raised to the protein, polypeptide or polynucleotide according to the invention, interferes with the immediate physical interaction between pathogen and mammalian host.

As used herein, the term "essential", when used in connection with a gene or gene product, means that the host cannot survive without, or is significantly growth compromised, in the absence or depletion of functional product. An "essential gene" is thus one that encodes a product that is beneficial, or preferably necessary, for cellular growth *in vitro* in a medium appropriate for growth of a strain having a wild-type allele corresponding to the particular gene in question. Therefore, if an essential gene is inactivated or inhibited, that cell will grow significantly more slowly than a wild-type strain or even not at all. Preferably, growth of a strain in which such a gene has been inactivated will be less than 20%, more preferably less than 10%, most preferably less than 5% of the growth rate of the wild-type, or the rate will be zero, in the growth medium. Preferably, in the absence of activity provided by a product of the gene, the cell will not grow at all or will be non-viable, at least under culture conditions similar to normal *in vivo* growth conditions. For example, absence of the biological activity of certain enzymes involved in bacterial cell wall synthesis can result in the lysis of cells under normal osmotic conditions, even though protoplasts can be maintained under controlled osmotic conditions. Preferably, but not necessarily, if such a

gene is inhibited, e.g., with an antibacterial agent or a phage product, the growth rate of the inhibited bacteria will be less than 50%, more preferably less than 30%, still more preferably less than 20%, and most preferably less than 10% of the growth rate of the uninhibited bacteria. As recognized by those skilled in the art, the degree of growth inhibition will generally depend upon the concentration of the inhibitory agent. In the context of the invention, essential genes are generally the preferred targets of antimicrobial agents. Essential genes can encode "target" molecules directly or can encode a product involved in the production, modification, or maintenance of a target molecule.

As used herein, "target" refers to a biomolecule or complex of biomolecules that can be acted on by an exogenous agent or compound, thereby modulating, preferably inhibiting, growth or viability of a bacterial cell. A target may be a nucleic acid sequence or molecule, or a polypeptide or a region of a polypeptide.

As used herein, the term "signal that is generated by interaction of a *S. aureus* polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or fragments thereof to a 3A ORF 33 or fragment thereof" or the like refers to the measurable indicator of polypeptide interaction in a binding assay, wherein the interacting polypeptide comprises the amino acid sequence of SEQ ID NO: 2, fragment thereof or variant thereof and 3A ORF 33, fragment thereof or variant thereof. As used herein, the term "signal that is generated by activation or inhibition of a *S. aureus* polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or fragments thereof" refers to the measurable indicator of polypeptide comprising the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof, activity in an assay of STAAU_R4 activity. For example, the signal may include, but is not limited to (i) a signal resulting from binding of 3A ORF 33 to a STAAU_R4 polypeptide, including a fluorescence signal (time-resolved fluorescence resonance energy transfer assay; fluorescence polarization assay), spectrophotometer absorbance measurement of a colourimetric signal (phage display ELISA), mass change measurement (surface

plasmon resonance analysis), or a viability measurement on selective medium (yeast two-hybrid analysis)

As used herein, the term "standard", used in reference to polypeptide activity, means the amount of activity observed or detected (directly or indirectly) in a given assay performed in the absence of a candidate compound. A "standard" serves as a reference to determine the effect, positive or negative, of a candidate compound on polypeptide activity.

As used herein, the term "increase in activity" refers to an enhanced level of measurable activity of a polypeptide in a given assay in the presence of a candidate compound relative to the measurable level of activity in the absence of a candidate compound. Activity is considered increased according to the invention if it is at least 10% greater, 20% greater, 50% greater, 75% greater, 100% greater or more, up to 2-fold, 5-fold, 10-fold, 20-fold, 50-fold, 100-fold or more than in the absence of a candidate compound.

As used herein, the term "decrease in activity" refers to a reduced level of measurable activity of a polypeptide in a given assay in the presence of a candidate compound relative to the measurable level of activity in the absence of a candidate compound. Activity is considered decreased according to the invention if it is at least 10% less, preferably 15% less, 20% less, 50% less, 75% less, or even 100% less (i.e., no activity) than that observed in the absence of a candidate compound.

As used herein, the term "conditions that permit their interaction", when used in reference to a *S. aureus* polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or fragments thereof, and a candidate compound means that the two entities are placed together, whether both in solution or with one immobilized or restricted in some way and the other in solution, wherein the parameters (e.g., salt, detergent, protein or candidate compound concentration, temperature, and redox potential, among others) of the solution are such that the *S. aureus* polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or fragments thereof, and the candidate compound

may physically associate. Conditions that permit protein:candidate interaction include, for example, the conditions described herein for TR-FRET, fluorescent polarization, Surface Plasmon Resonance and Phage display assays.

As used herein, the term "detectable change in a measurable parameter of STAAU_R4" refers to an alteration in a quantifiable characteristic of a *S. aureus* STAAU_R4 polypeptide.

As used herein, the term "agonist" refers to an agent or compound that enhances or increases the activity of a *S. aureus* STAAU_R4 polypeptide or polynucleotide. An agonist may be directly active on a *S. aureus* STAAU_R4 polypeptide or polynucleotide, or it may be active on one or more constituents in a pathway that leads to enhanced or increased activity of a *S. aureus* STAAU_R4 polypeptide or polynucleotide.

As used herein, the term "antagonist" refers to an agent or compound that reduces or decreases the activity of a *S. aureus* STAAU_R4 polypeptide or polynucleotide. An antagonist may be directly active on a *S. aureus* STAAU_R4 polypeptide or polynucleotide, or it may be active on one or more constituents in a pathway that leads to reduced or decreased activity of a *S. aureus* STAAU_R4 polypeptide or polynucleotide.

As used herein, the term "antibacterial agent" or "antibacterial compound" refers to an agent or compound that has a bacteriocidal or bacteriostatic effect on one or more bacterial strains, preferably such an agent or compound is bacteriocidal or bacteriostatic on at least *S. aureus*.

As used herein, the term "synthesizing" refers to a process of chemically synthesizing a compound.

As used in the context of treating a bacterial infection a "therapeutically effective amount", "pharmaceutically effective amount" or "amount sufficient to provide a therapeutic effect" indicates an amount of an antibacterial agent which has a therapeutic effect. This generally refers to the inhibition, to some extent, of the normal cellular functioning of bacterial cells required for continued bacterial infection. Further, as used herein, a therapeutically effective

amount means an amount of an antibacterial agent that produces the desired therapeutic effect as judged for example by clinical trial results and/or animal models. This amount can be routinely determined by one skilled in the art and will vary depending on several factors, such as the particular bacterial strain involved and the particular antibacterial agent used. In the same context, an "amount sufficient to reduce adhesion" of a bacterium to a tissue or tissue surface indicates an amount of an antibacterial agent that is effective for prophylactically preventing or reducing the extent of bacterial infection of the given tissue or tissue surface. The same principle applies to the terminology "prophylactically effective amount".

As used in the context of treating a bacterial infection, contacting or administering the antimicrobial agent 'in combination with existing antimicrobial agents' refer to a concurrent contacting or administration of the active compound with antibiotics to provide a bactericidal or growth inhibitory effects beyond the individual bactericidal or growth inhibitory effects of the active compound or the antibiotic. Existing antibiotic refers for example to the group consisting of penicillins, cephalosporins, imipenem, monobactams, aminoglycosides, tetracyclines, sulfonamides, trimethoprim/sulfonamide, fluoroquinolones, macrolides, vancomycin, polymyxins, chloramphenicol and lincosamides.

As used herein, a "tissue" refers to an aggregation of cells of one or more cell types which together perform one or more specific functions in an organism. As used herein, a "tissue surface" refers to that portion of a tissue that forms a boundary between a given tissue and other tissues or the surroundings of the tissue. A tissue surface may refer to an external surface of an animal, for example the skin or cornea, or, alternatively, the term may refer to a surface that is either internal, for example, the lining of the gut, or to a surface that is exposed to the outside surroundings of the animal only as the result of an injury or a surgical procedure.

As used herein, the term "measuring the binding of a candidate compound" refers to the use of an assay preferably permitting the quantitation of the amount of a candidate compound physically associated with a *S. aureus* STAAU_R4 polypeptide, fragment or variant thereof.

A "candidate compound" as used herein, is any compound with a potential to modulate the expression or activity of a *S. aureus* STAAU_R4 polypeptide.

As used herein, the term "directly or indirectly detectably labeled" refers to the attachment of a moiety to a candidate compound that renders the candidate compound either directly detectable (e.g., an isotope or a fluorophore) or indirectly detectable (e.g., an enzyme activity, allowing detection in the presence of an appropriate substrate, or a specific antigen or other marker allowing detection by addition of an antibody or other specific indicator).

A "method of screening" refers to a method for evaluating a relevant activity or property of a large plurality of compounds, rather than just one or a few compounds. For example, a method of screening can be used to conveniently test at least 100, more preferably at least 1000, still more preferably at least 10,000, and most preferably at least 100,000 different compounds, or even more. In a particular embodiment, the method is amenable to automated, cost-effective high throughput screening on libraries of compounds for lead development.

In a related aspect or in preferred embodiments, the invention provides a method of screening for potential antibacterial agents by determining whether any of a plurality of compounds, preferably a plurality of small molecules, is active on STAAU_R4. Preferred embodiments include those described for the above aspect, including embodiments which involve determining whether one or more test compounds bind to or reduce the level of activity of a bacterial target, and embodiments which utilize a plurality of different targets as described above.

As used herein, the term "simultaneously" when used in connection with the assays of the present invention, refers to the fact that the

specified components or actions at least overlap in time, and is thus not restricted to the fact that the initiation and termination points are identical. For certainty, a simultaneous contact of a STAAU_R4 polypeptide with a candidate compound and a bacteriophage polypeptide, for example, is an overlap in contact periods, which can, but does not necessarily reflect the fact that the latter two are introduced into an assay mixture at the exact same time.

The term "compounds" preferably includes, but is not limited to, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention, such as for example inhibitory ORF gene product or target thereof, and thereby inhibit, extinguish or enhance its activity or expression. Potential compounds may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same site(s) on a binding molecule, such as a bacteriophage gene product, thereby preventing bacteriophage gene product from binding to STAAU_R4 polypeptides.

The term "compounds" is also meant to include small molecules that bind to and occupy the binding site of a polypeptide, thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Preferred potential compounds include compounds related to and variants of inhibitory ORF encoded by a bacteriophage and of STAAU_R4 and any homologues and/or peptido-mimetics and/or fragments thereof. Other examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented. Other potential compounds include antisense molecules (see Okano, 1991 J. Neurochem. 56, 560; see also "Oligodeoxynucleotides as

Antisense Inhibitors of Gene Expression", CRC Press, Boca Raton, FL (1988), for a description of these molecules).

As used herein, the term "library" refers to a collection of 100 compounds, preferably of 1000, still more preferably 5000, still more preferably 10,000 or more, and most preferably of 50,000 or more compounds.

As used herein, the term "small molecule" refers to compounds having molecular mass of less than 3000 Daltons, preferably less than 2000 or 1500, still more preferably less than 1000, and most preferably less than 600 Daltons. Preferably but not necessarily, a small molecule is not an oligopeptide.

As used herein, the term "mimetic" refers to a compound that can be natural, synthetic, or chimeric and is structurally and functionally related to a reference compound. In terms of the present invention, a "peptidomimetic," for example, is a non-peptide compound that mimics the activity-related aspects of the 3-dimensional structure of a peptide or polypeptide, for example a compound that mimics the structure of a peptide or active portion of a phage- or bacterial ORF-encoded polypeptide.

As used herein, the term "bacteriophage inhibitor protein" refers to a protein encoded by a bacteriophage nucleic acid sequence, which inhibits bacterial function in a host bacterium. Thus, it is a bacteria-inhibiting phage product. The term "bacteriophage inhibitor protein" encompasses a fragment, derivative, or active portion of a bacteriophage inhibitor protein.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either STAAU_R4 or its target molecule or ligand to facilitate separation of complexed from uncomplexed forms of one or both of the proteins or polypeptides, as well as to accommodate automation of the assay. Binding of a test compound to a STAAU_R4 protein (or fragment, or variant thereof) or interaction of a STAAU_R4 protein with a target molecule or ligand in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants.

Examples of such vessels include microtitre plates, test tubes and micro-centrifuge tubes.

In one embodiment a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase (GST)/STAAU_R4 fusion proteins or GST/target fusion proteins (e.g. GST/3A ORF 33) can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or STAAU_R4 protein and the mixture incubated under conditions conducive to complex formation (e.g. at physiological conditions for salt and pH). Following incubation the beads or microtitre plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of STAAU_R4 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices (and well-known in the art) can also be used in the screening assays of the invention. For example, either a STAAU_R4 protein or a STAAU_R4 target molecule or ligand can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated STAAU_R4 protein or target molecules or ligand can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with target molecules or ligand but which do not interfere with binding of the STAAU_R4 protein (or part thereof) to its target molecule or ligand can be derivatized to the wells of the plate, and unbound target or STAAU_R4 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with

the STAAU_R4 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the STAAU_R4 protein.

As used herein, the term "active portion", when refering to a bacteriophage-derived sequence, relates to an epitope, a catalytic or regulatory domain, or a fragment of a bacteriophage inhibitor protein that is responsible for, or a significant factor in, bacterial target inhibition. The active portion preferably may be removed from its contiguous sequences and, in isolation, still effect inhibition.

As used herein, the term "treating a bacterial infection" refers to a process whereby the growth and/or metabolic activity of a bacterium or bacterial population in a host, preferably a mammal, more preferably a human, is inhibited or ablated. Similarly, the term "preventing a bacterial infection" refers to a process whereby the growth and/or metabolic activity of a bacterium or bacterial population in a host, preferably a mammal, more preferably a human, at risk of being infected is inhibited or ablated.

As used herein, the term "bacterium" refers to a single bacterial strain and includes a single cell and a plurality or population of cells of that strain unless clearly indicated to the contrary. In reference to bacteria or bacteriophage, the term "strain" refers to bacteria or phage having a particular genetic content. The genetic content includes genomic content as well as recombinant vectors. Thus, for example, two otherwise identical bacterial cells would represent different strains if each contained a vector, *e.g.*, a plasmid, with different inserts.

As used herein, the term "diagnosing" refers to the identification of an organism or strain of an organism responsible for a bacterial infection.

As used herein, the term "infection with *Staphylococcus aureus*" refers to the presence, growth or proliferation of cells of a *S. aureus* strain within, or on a surface of, an animal, such as a mammal, preferably a human.

As used herein, the term "bacteriophage 3A ORF 33 -encoded polypeptide" refers to a polypeptide encoded by SEQ ID NO: 3 or to a fragment or derivative thereof encompassing an active portion of a bacteriophage 3A ORF 33 -encoded polypeptide of sequence disclosed in SEQ ID NO: 4.

As used herein, the term "polypeptide complex" refers to a combination of two or more polypeptides in a physical association with each other. It is preferred that such a physical association be required for some aspect of the activity of one or more of the polypeptides in such a polypeptide complex.

As used herein, the term "physical association" refers to an interaction between two moieties involving contact between the two moieties.

As used herein, the term "bodily material(s)" means any material derived from an individual or from an organism infecting, infesting or inhabiting an individual, including but not limited to, cells, tissues and waste, such as, bone, blood, serum, cerebrospinal fluid, semen, saliva, muscle, cartilage, organ tissue, skin, urine, stool or autopsy materials.

As used herein, the term "disease(s)" means any disease caused by or related to infection by a bacterium, including, for example, otitis media, conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural empyema and endocarditis, and most particularly meningitis, such as for example infection of cerebrospinal fluid.

As used herein, the term "fusion protein(s)" refers to a protein encoded by a gene comprising amino acid coding sequences from two or more separate proteins fused in frame such that the protein comprises fused amino acid sequences from the separate proteins.

As used herein, the term "host cell(s)" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

As used herein, the term "immunologically equivalent derivative(s)" encompasses a polypeptide, polynucleotide, or the equivalent of either which when used in a suitable formulation to raise antibodies in a

vertebrate, results in antibodies that act to interfere with the immediate physical interaction between pathogen and mammalian host.

As used herein, the term "immunospecific" means that characteristic of an antibody whereby it possesses substantially greater affinity for the polypeptides of the invention or the polynucleotides of the invention than its affinity for other related polypeptides or polynucleotides respectively, particularly those polypeptides and polynucleotides in the prior art.

As used herein, the term "individual(s)" means a multicellular eukaryote, including, but not limited to a metazoan, a mammal, an ovid, a bovid, a simian, a primate, and a human.

As used herein, the term "Organism(s)" means a (i) prokaryote, including but not limited to, a member of the genus *Streptococcus, Staphylococcus, Bordetella, Corynebacterium, Mycobacterium, Neisseria, Haemophilus, Actinomycetes, Streptomycetes, Nocardia, Enterobacter, Yersinia, Fancisella, Pasturella, Moraxella, Acinetobacter, Erysipelothrix, Branhamella, Actinobacillus, Streptobacillus, Listeria, Calymmatobacterium, Brucella, Bacillus, Clostridium, Treponema, Escherichia, Salmonella, Kleibsiella, Vibrio, Proteus, Erwinia, Borrelia, Leptospira, Spirillum, Campylobacter, Shigella, Legionella, Pseudomonas, Aeromonas, Rickettsia, Chlamydia, Borrelia and Mycoplasma, and further including, but not limited to, a member of the species or group, Group A Streptococcus, Group B Streptococcus, Group C Streptococcus, Group D Streptococcus, Group G Streptococcus, Streptococcus pneumoniae, Streptococcus pyogenes, Streptococcus agalactiae, Streptococcus faecalis, Streptococcus faecium, Streptococcus durans, Neisseria gonorrheae, Neisseria meningitidis, Staphylococcus aureus, Staphylococcus epidermidis, Corynebacterium diptheriae, Gardnerella vaginalis, Mycobacterium tuberculosis, Mycobacterium bovis, Mycobacterium ulcerans, Mycobacterium leprae, Actinomyctes israelii, Listeria monocytogenes, Bordetella pertusis, Bordatella parapertusis, Bordetella bronchiseptica, Escherichia coli, Shigella dysenteriae, Haemophilus influenzae, Haemophilus aegyptius, Haemophilus parainfluenzae,*

Haemophilus ducreyi, Bordetella, Salmonella typhi, Citrobacter freundii, Proteus mirabilis, Proteus vulgaris, Yersinia pestis, Kleibsiella pneumoniae, Serratia marcessens, Serratia liquefaciens, Vibrio cholera, Shigella dysenterii, Shigella flexneri, Pseudomonas aeruginosa, Franscisella tularensis, Brucella abortis, Bacillus anthracis, Bacillus cereus, Clostridium perfringens, Clostridium tetani, Clostridium botulinum, Treponema pallidum, Rickettsia rickettsii and *Chlamydia trachomitis*, (ii) an archaeon, including but not limited to *Archaebacter*, and (iii) a unicellular or filamenous eukaryote, including but not limited to, a protozoan, a fungus, a member of the genus Saccharomyces, *Kluveromyces*, or *Candida*, and a member of the species *Saccharomyces ceriviseae, Kluveromyces lactis*, or *Candida albicans*.

As used herein, the term "recombinant expression system(s)" refers to a system in which vectors comprising sequences encoding polypeptides of the invention or portions thereof, or polynucleotides of the invention are introduced or transformed into a host cell or host cell lysate for the production of the polynucleotides and polypeptides of the invention.

As used herein, the term "artificially synthesized" when used in reference to a peptide, polypeptide or polynucleotide means that the amino acid or nucleotide subunits were chemically joined *in vitro* without the use of cells or polymerizing enzymes. The chemistry of polynucleotide and peptide synthesis is well known in the art.

In addition to the standard single and triple letter representations for amino acids, the term "X" or "Xaa" may also be used in describing certain polypeptides of the invention. "X" and "Xaa" mean that any of the twenty naturally occurring amino acids may appear at such a designated position in the polypeptide sequence.

As used herein, the term "specifically binding" in the context of the interaction of two polypeptides means that the two polypeptides physically interact via discrete regions or domains on the polypeptides, wherein the interaction is dependent upon the amino acid sequences of the interacting

domains. Generally, the equilibrium binding concentration of a polypeptide that specifically binds another is in the range of about 1 mM or lower, more preferably 1 uM or lower, preferably 100 nM or lower, 10 nM or lower, 1 nM or lower, 100 pM or lower, and even 10 pM or lower.

As used herein, the term "decrease in the binding" refers to a drop in the signal that is generated by the physical association between two polypeptides under one set of conditions relative to the signal under another set of reference conditions. The signal is decreased if it is at least 10% lower than the level under reference conditions, and preferably 20%, 40%, 50%, 75%, 90%, 95% or even as much as 100% lower (i.e., no detectable interaction).

As used herein, the term "detectable marker", when used in the context of a yeast two-hybrid assay, refers to a polypeptide that confers a trait upon a cell expressing that polypeptide that signals the presence or amount of that polypeptide expressed. Detectable markers are encoded on plasmids that may exist episomally or may be integrated into the genome of a host cell. Detectable markers include, but are not limited to, polypeptides encoding enzymes allowing colorimetric or fluorescent detection (e.g., *E. coli LacZ*, which catalyzes the conversion of the substrate analog X-gal to generate a blue color), polypeptides encoding enzymes conferring antibiotic resistance, and polypeptides encoding enzymes conferring the ability of a yeast strain to grow on medium lacking a given component (i.e., critical for the relief of auxotrophy).

As used herein, the term "results in the expression of a detectable marker" means that the interaction of factors necessary to permit the expression of a detectable marker (e.g., two-hybrid transactivation domain and DNA binding domain fusion proteins) causes the transactivation and translation of detectable levels of a detectable marker. A "detectable level" is that level of expression that can be differentiated from background expression occurring in the substantial absence of one or more factors or conditions necessary for marker expression. Detectable levels will vary depending upon the nature of the

detectable marker, but will generally consist of levels at least about 10% or more greater than the background level of a given marker.

As used herein, the term "decrease in the expression" refers to a drop in the expression of a detectable marker under one set of conditions relative to the expression under another set of reference conditions. The expression of a detectable marker is decreased if it is at least 10% lower than the level under reference conditions, and preferably 20%, 40%, 50%, 75%, 90%, 95% or even as much as 100% lower (i.e., not expressed).

Identification of the *S. aureus* STAAU_R4 sequence

The methodology used to identify the STAAU_R4 polypeptide is described in detail in U.S. Provisional Patent Application 60/110,992, filed December 3, 1998, and PCT International Application WO1999/IB99/02040, filed December 3, 1999. Briefly, this PCT application concerns bacteriophages that can infect a selected bacterium. The sequencing and characterization of the phage genetic information allow the identification of all open reading frames (ORFs) encoded by the phage, including those that are essential or instrumental in inhibiting their host. Each ORF is identified using computer softwares and individually expressed in the host. The effect of this expression on host viability is then measured. Identification of ORFs from the phage genome which inhibit the host bacterium both provides a compound that could be used as a bacterial inhibitor compound *per se* (or derivatized or modified to obtain further inhibitors) and as a tool for the identification of the bacterial target affected by the phage-encoded inhibitor.

Using methodology described in detail in Example 1 and 2, a *S. aureus* polypeptide that specifically bound the bacterial growth inhibitory 3A phage ORF 33 protein was isolated. Briefly, the 3A ORF 33 protein was used as a ligand in an affinity chromatography binding step with *S. aureus* protein extract. The selected *S. aureus* interacting polypeptide was purified and further analyzed by tryptic digestion and mass spectrometry using MALDI-ToF technology [Qin, J., *et al.* (1997) *Anal. Chem.* 69:3995-4001]. Computational analysis

(http://prowl.rockfeller.edu/cgi-bin/ProFound) of the mass spectrum obtained identifies the corresponding ORF in the *S. aureus* nucleotide sequence in the University of Oklahoma *S. aureus* genomic database at http://www.genome.ou.edu/staph.html. The interaction between 3A ORF 33 and the candidate target protein, herein referred as STAAU_R4, was also confirmed in affinity blotting assays. Comparison of the ORF of the *S. aureus* contig that encodes a tryptic peptide similar to that identified in the *S. aureus* phage 3A ORF 33 binding studies with all other sequences in the public domain databases revealed that STAAU_R4 is related to PlsX, a protein implicated in fatty acid/phospholipid synthesis (Fig. 8). As shown in Fig. 9A, the degree of relatedness of the identified *B. subtilis* ORF to the STAAU_R4 protein is 53% identity and 71% similarity at the amino acid level across the entire sequence.

Function of PlsX (STAAU_R4)

Fatty acid biosynthesis is essential to the viability of all cells because it yields a wide variety of lipid molecules some of which are key structural components of membranes and others which have important roles in signalling and in energy storage. The importance of phospholipids, for example, is underscored by the lack of genetic defects in the metabolism of these lipids in humans. In bacteria, lipid metabolism is similarly of critical importance to the cell. In *E. coli*, eight enzyme-catalyzed reactions convert acetyl-coenzyme A into fatty acids. Three subsequent reactions append acyl chains onto glycerol-3-phosphate to yield CDP diacylglycerol, a key branchpoint molecule for the biosynthesis of important cellular phospholipids including phosphatidylserine, phosphatidylethanolamine, phosphatidylglycerol, and cardiolipin.

The initial reaction of phospholipid biosynthesis in *E. coli* is catalyzed by the product of the *plsB* gene, glycerol-3-phosphate acyltransferase. This enzyme preferentially utilizes saturated fatty acyl derivatives (palmitoyl-coenzymeA or palmitoyl-acyl carrier protein) for the initial acylation of glycerol-3-phosphate. Interestingly, the glycerol-3-phosphate auxotrophic mutant of *E. coli* that was originally described by Bell [Bell, R.M. (1974) J. Bacteriol. 117: 1065-

1076] was found to contain mutations at two loci; both mutations were necessary to observe the dependence of the strain on glycerol-3-phosphate for growth [Larson, T.J., Ludtke, D.N., and Bell, R.M. (1984) J. Bacteriol. 160: 711-717]. Independent transposon insertions from wild-type cells, mediated by phage P1 cotransduction, were sought which could correct the auxotrophy of the mutant strain. As expected, a portion of the transposon insertions were cotransducible with markers in the *plsB* region. However, the remainder mapped to a locus termed *plsX* in a distinct region of the chromosome. Indeed, both *plsB26* and *plsX50* mutations are necessary and sufficient for conferral of glycerol-3-phosphate auxotrophy.

Wild-type PlsX may function either to elevate the intracellular levels of glycerol-3-phosphate or to increase the affinity of the defective glycerol-3-phosphate acyltransferase, encoded by *plsB26*, for its substrate, allowing the mutant PlsB to function [Larson, T.J., Ludtke, D.N., and Bell, R.M. (1984) J. Bacteriol. 160: 711-717]. PlsX apparently does not influence glycerol-3-phosphate acyltransferase activity *in vitro* since the V_{max} and the apparent K_m defects resulting from the *plsB26* mutation were independent of the presence of wild type PlsX. However, restoration of a wild-type copy of PlsX to a strain of *E. coli* bearing the *plsB26 plsX50* mutations and a temperature-sensitive mutation in PlsC, encoding the 1-acyl glycerol-3-phosphate acyltransferase, was not sufficient to restore lysophosphatidic acid synthesis *in vivo* at the nonpermissive temperature [Heath, R.J., Goldfine, H., and Rock, C.O. (1997) J. Bacteriol. 179: 7257-7263]. Thus, PlsX cannot restore glycerol-3-phosphate activity *in vivo*.

Functional homologues to *E. coli plsX* exist in several bacterial species, including *Salmonella typhimurium* [Zhang, Y., and Cronan Jr., J.E. (1998) J. Bacteriol. 180: 3295-3303] and the Gram-positive bacteria *Clostridium butyricum* [Heath, R.J., Goldfine, H., and Rock, C.O. (1997) J. Bacteriol. 179: 7257-7263] and *Bacillus subtilis* [Morbidoni, H., de Mendoza, D., and Cronan, Jr., J.E. (1996) J. Bacteriol. 178: 4794-4800]. To date, no homologue of the *S. aureus* PlsX has been described or isolated. In Gram-positive bacteria, while the function

of PlsX remains to be determined, *plsX* is likely to be an essential gene based on indications that repeated attempts to disrupt the gene were unsuccessful [Morbidoni, H., de Mendoza, D., and Cronan, Jr., J.E. (1996) J. Bacteriol. 178: 4794-4800].

The demonstration that bacteriophage have adapted to inhibiting a host bacterium by acting on a particular cellular component or target provides a strong indication that that component is an appropriate target for developing and using antibacterial agents, e.g. in therapeutic treatments. The present invention provides additional guidance over mere identification of bacterial essential genes, as the present invention also provides an indication of accessibility of the target to an inhibitor, and an indication that the target is sufficiently stable over time (e.g., not subject to high rates of mutation) as phage acting on that target were able to develop and persist. Thus the present invention identifies STAAU_R4 as particularly likely to bean appropriate target for development of antibacterial agents.

Identification of the surface of interaction on STAAU_R4

This invention relates, in part, to a specific interaction between a growth-inhibitory protein encoded by the *S. aureus* bacteriophage genome and an essential *S. aureus* protein. In one embodiment, this interaction forms the basis for drug screening assays. More specifically, the invention relates to the interacting domains of the protein encoded by the *S. aureus* STAAU_R4 and the *S. aureus* bacteriophage 3A ORF 33 proteins, forming the basis for screening assays. The invention provides a method for the identification of 3A ORF 33 and, more preferably, STAAU_R4 polypeptide fragments which are involved in the interaction between STAAU_R4 and 3A ORF 33.

Several approaches and techniques known to those skilled in the art can be used to identify and to characterize interacting fragments of STAAU_R4 and 3A ORF 33. These fragments may include, for example, truncation polypeptides having a portion of an amino acid sequence of any of the

two proteins, or variants thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence.

Fragments of STAAU_R4 and 3A ORF 33 can be cloned by genetic recombinant technology and tested for interaction using a yeast two-hybrid assay as exemplified below.

Partial proteolysis of proteins in solution is one method to delineate the domain boundaries in multi-domain proteins. By subjecting proteins to limited digestion, the most accessible cleavage sites are preferentially hydrolyzed. These cleavage sites preferentially reside in less structured regions which include loops and highly mobile areas typical of the joining amino acids between highly structures domains. Purified STAAU_R4 and 3A ORF 33 proteins can be subjected to partial proteolysis. The proteolysis can be performed with low concentrations of proteases (trypsin, chymotrypsin, endoproteinase Glu-C, and Asp-N) with STAAU_R4 or 3A ORF 33 in solution, resulting in the generation of defined proteolytic products as observed by SDS-PAGE. An acceptable concentration and reaction time is defined by the near complete conversion of the full-length protein to stable proteolytic products. The proteolytic products are then subjected to affinity chromatography containing the appropriated partner of interaction (3A ORF 33 or STAAU_R4 purified proteins) to determine a protein sub-region able to interact. Interacting domains are identified by mass spectrometry to determine both the intact fragment mass and the completely digested with trypsin (by in-gel digestion) to better determine the amino acid residues contained within the partial proteolytic fragment. Using both sets of data, the amino acid sequence of the partial proteolytic fragment can be precisely determined.

Another approach is based on peptide screening using different portions of 3A ORF 33 or STAAU_R4 to identify minimal peptides from each polypeptide that are able to disrupt the interaction between the two proteins. It is assumed that fragments able to prevent interaction between STAAU_R4 and 3A ORF 33 correspond to domains of interaction located on either of the two

interacting proteins. The different peptide fragments can be screened as competitors of interaction in protein: protein binding assays such as the ones described below. Fine mapping of interaction site(s) within a protein can be performed by an extensive screen of small overlapping fragments or peptides spanning the entire amino acid sequence of the protein.

Suitable STAAU_R4 and 3A ORF 33 -derived amino acid fragments representative of the complete sequence of both proteins can be chemical synthesis. For instance, in the multipin approach, peptides are simultaneously synthesis by the assembly of small quantities of peptides (ca. 50 nmol) on plastic pins derivatized with an ester linker based on glycolate and 4-(hydroxymethyl) benzoate [Maeji 1991 Pept Res, 4:142-6].

S. aureus STAAU_R4 polypeptides

In one aspect of the invention there are provided polypeptides of *S. aureus* referred to herein as "STAAU_R4" and "STAAU_R4 polypeptides" as well as biologically, diagnostically, prophylactically, clinically or therapeutically useful variants thereof, and compositions comprising the same.

Among the particularly preferred embodiments of the invention are variants of *S. aureus* STAAU_R4 polypeptides encoded by naturally occurring alleles of the STAAU_R4 gene. The present invention provides for an isolated polypeptide which comprises or consists of: (a) an amino acid sequence which has at least 50% identity, preferably at least 55% identity, preferably at least 60% identity, preferably at least 70% identity, preferably at least 80% identity, more preferably at least 90%, yet more preferably at least 95%, most preferably at least 97-99%, or exact identity, over the entire length of SEQ ID NO: 2; or b) an amino acid sequence that has at least 70% similarity, at least 75% similarity, at least 80% similarity, at least 90% similarity, at least 95% similarity, at least 97-99% similarity or even 100% similarity over the entire length of SEQ ID NO: 2.

The polypeptides of the invention include a polypeptide of Fig. 1 (SEQ ID NO: 2) (in particular the mature polypeptide) as well as polypeptides and fragments, particularly those which have a biological activity of STAAU_R4,

and also those which have at least 50% identity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2 or the relevant portion, preferably at least 55%, 60%, 70%, or 80% identity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2, more preferably at least 90% identity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2 and still more preferably at least 95% identity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2 and yet still more preferably at least 99% identity or exact identity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2.

The polypeptides of the invention also include a polypeptide or protein fragment that has at least 70%, 75%, 80% or 90% similarity, 95% similarity or even 97-99% similarity over 50 or more amino acids to a polypeptide of SEQ ID NO: 2.

It is most preferred that a polypeptide of the invention is derived from *S. aureus*, however, it may be obtained from other organisms of the same taxonomic genus. A polypeptide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Fragments of STAAU_R4 also are included in the invention. These fragments may include, for example, truncation polypeptides having a portion of an amino acid sequence of Fig. 1 (SEQ ID NO: 2), a fragment or a variant thereof, such as a continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence. Degradation forms of the polypeptides of the invention produced by or in a host cell, particularly *S. aureus*, are also preferred. Further preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix-forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Fragments of STAAU_R4 may be expressed as fusion proteins with other proteins or protein fragments.

Preferred fragments also include an isolated polypeptide comprising an amino acid sequence having at least 10, 15, 20, 30, 39, 50, 100, 200, 250, 300 or or more contiguous amino acids from the amino acid sequence of SEQ ID NO: 2.

Also preferred are biologically "active" fragments which are those fragments that mediate activities of *S. aureus* STAAU_R4, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those fragments that are antigenic or immunogenic in an animal, especially in a human. Particularly preferred are fragments comprising domains that confer a function essential for viability of *S. aureus*.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention.

S. aureus Polynucleotides

It is an object of the invention to provide polynucleotides that encode STAAU_R4 polypeptides, particularly polynucleotides that encode the polypeptide herein designated *S. aureus* STAAU_R4.

In one aspect of the invention, a polynucleotide is provided that comprises a region encoding a *S. aureus* STAAU_R4 polypeptide, the polynucleotide comprising a sequence set out in SEQ ID NO: 1. Such a polynucleotide encodes a full length STAAU_R4 gene, a fragment or a variant thereof. It is contemplated that this full-length gene is essential to the growth and/or survival of an organism which possesses it, such as *S. aureus*.

As a further aspect of the invention there are provided isolated nucleic acid molecules encoding and/or expressing a fragment of a full-length STAAU_R4 polypeptide, particularly a *S. aureus* STAAU_R4 polypeptide, a fragment or a variant thereof. Further embodiments of the invention include biologically, diagnostically, prophylactically, clinically or therapeutically useful

polynucleotides, polypeptides, variants thereof, and compositions comprising same.

A polynucleotide of the invention is obtained using *S. aureus* cells as starting material, the nucleotide sequence information disclosed in SEQ ID NO:1, and standard cloning and screening methods, such as those for cloning and sequencing chromosomal DNA fragments from bacteria. For example, to obtain a polynucleotide sequence of the invention, such as the polynucleotide sequence disclosed as in SEQ ID NO: 1, a library of clones of chromosomal DNA of *S. aureus* in *E. coli* or another suitable host is probed with a radiolabeled oligonucleotide, preferably a 17-mer or longer, derived from a partial sequence. Clones carrying DNA identical to that of the probe can be distinguished using stringent hybridization conditions. As herein used, the terms "stringent conditions" and "stringent hybridization conditions" mean hybridization occurring only if there is at least 95% and preferably at least 97% identity between the sequences. A specific example of stringent hybridization conditions is of an overnight incubation of a hybridization support (e.g., a nylon or nitrocellulose membrane) at 42°C in a solution comprising: 1 X 10^6 cpm/ml labeled probe, 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/ml of denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at 65°C. Hybridization and wash conditions are well known to those skilled in the art and are exemplified in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), particularly Chapter 11 therein. Solution hybridization may also be used with the polynucleotide sequences provided by the invention. By sequencing the individual clones thus identified by hybridization, it is possible to confirm the identity of the clone.

Alternatively, an amplification process can be utilized to isolate the polynucleotide. In this approach, the sequence disclosed as SEQ ID NO:1 is targeted by two oligonucleotides, one identical to a sequence on the coding DNA

strand at or upstream of the ATG initiation codon and the other which anneals to the opposite strand at or downstream of the stop codon. Priming from these oligonucleotides in a polymerase chain reaction yields a full-length gene coding sequence. Such suitable techniques are described by Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

In a further aspect, the present invention provides for an isolated polynucleotide comprising or consisting of: (a) a polynucleotide sequence which has at least 60% identity, preferably at least 70% identity, more preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95%, most preferably at least 97-99% or exact identity, to that of SEQ ID NO: 1; (b) a polynucleotide sequence encoding a polypeptide which has at least 50% identity, preferably at least 55% identity, preferably at least 60% identity, more preferably at least 70% identity, more preferably at least 80% identity, more preferably at least 90%, yet more preferably at least 95%, most preferably at least 97-99% or exact identity to SEQ ID NO: 2 over the entire length of SEQ ID NO: 2; or the complement of a sequence of (a) or (b) above.

The invention provides a polynucleotide sequence identical over its entire length to the coding sequence of SEQ ID NO:1. Also provided by the invention is a coding sequence for a mature polypeptide or a fragment thereof by itself as well as a coding sequence for a mature polypeptide or a fragment in reading frame with another coding sequence, such as a sequence encoding a leader or secretory sequence, a pre-, or pro-, or prepro-protein sequence. The polynucleotide of the invention may also contain at least one non-coding sequence, including for example, but not limited to at least one non-coding 5' and 3' sequence, such as the transcribed but non-translated sequences, termination signals (such as rho-dependent and rho-independent termination signals), ribosome binding sites, Kozak sequences, sequences that stabilize or destabilize mRNAs, introns, and polyadenylation signals. The polynucleotide sequence may also comprise additional coding sequence encoding additional amino acids. For

example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain embodiments of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc. Natl. Acad. Sci. 86: 821-824 (1989), or an HA peptide tag [Wilson *et al.*, Cell 37: 767 (1984)], both of which may be useful in purifying polypeptide sequences fused to them. Polynucleotides of the invention also include, but are not limited to, polynucleotides comprising a structural gene and its naturally associated sequences that control gene expression.

While it is most preferred that a polynucleotide of the invention be derived from *S. aureus*, it may also be obtained from other organisms of the same taxonomic genus. A polynucleotide of the invention may also be obtained, for example, from organisms of the same taxonomic family or order.

Further preferred embodiments are polynucleotides encoding *S. aureus* STAAU_R4 variants that have the amino acid sequence of *S. aureus* STAAU_R4 polypeptide of SEQ ID NO: 2 in which several, a few, 10 to 50, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, modified, deleted and/or added, in any combination. Especially preferred among these polynucleotides are those encoding silent nucleotide alterations that do not alter the coding sequence or activities of *S. aureus* STAAU_R4 polypeptides they encode.

Preferred embodiments are polynucleotides encoding polypeptides that retain substantially the same biological function or activity as the mature polypeptide encoded by a DNA of SEQ ID NO: 1.

In accordance with certain preferred embodiments of this invention there are provided polynucleotides that hybridize, particularly under stringent conditions, to *S. aureus* STAAU_R4 polynucleotide sequences, such as those polynucleotides in Fig. 1.

The polynucleotides of the invention are useful as hybridization probes for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding genes that have a high degree of sequence

identity to the STAAU_R4 gene. Such probes generally will comprise at least 15 to about 100 residues or base pairs, although such probes will preferably have about 20 to 50 nucleotide residues or base pairs. Particularly preferred probes are about 20 to about 30 nucleotide residues or base pairs in length.

A coding region of a related STAAU_R4 gene from a bacterial species other than *S. aureus* may be isolated by screening a library using a DNA sequence provided in SEQ ID NO: 1 to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine to which member(s) of the library the probe hybridizes.

There are several methods available and well known to those skilled in the art to obtain full-length DNAs, or extend short DNAs, for example those based on the method of Rapid Amplification of cDNA Ends (RACE) [see, for example, Frohman, *et al.*, Proc. Natl. Acad. Sci. USA 85:8998-9002, 1988]. Recent modifications of the technique, exemplified by the MARATHON™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the MARATHON technology, cDNAs are prepared from mRNA extracted from a chosen cell and an 'adaptor' sequence is ligated onto each end. Nucleic acid amplification by PCR is then carried out to amplify the "missing" 5' end of the DNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using "nested" primers, that is, primers designed to anneal within the amplified product (typically an adaptor-specific primer that anneals further 3' in the adaptor sequence and a gene-specific primer that anneals further 5' in the selected gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length DNA constructed either by joining the product directly to the existing DNA to give a complete sequence, or by carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

The polynucleotides and polypeptides of the invention may be employed, for example, as research reagents and materials for discovery of treatments of and diagnostics for diseases, particularly human diseases, as further discussed herein relating to polynucleotide assays.

The polynucleotides of the invention that are oligonucleotides derived from a sequence of SEQ ID NO:1 are useful for the design of PCR primers in reactions to determine whether or not the polynucleotides identified herein in whole or in part are transcribed in bacteria in infected tissue. That is, the polynucleotides of the invention are useful for diagnosis of infection with a bacterial strain carrying those sequences. It is recognized that such sequences also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained.

The invention also provides polynucleotides that encode a polypeptide that is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide. Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in vivo*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having a mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

A polynucleotide of the invention thus may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences that are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more

prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

In addition to the standard A, G, C, T/U representations for nucleotides, the term "N" may also be used in describing certain polynucleotides of the invention. "N" means that any of the four DNA or RNA nucleotides may appear at such a designated position in the DNA or RNA sequence, except it is preferred that N is not a nucleotide that when taken in combination with adjacent nucleotide positions, read in the correct reading frame, would have the effect of generating a premature termination codon in such reading frame.

For each and every polynucleotide of the invention there is also provided a polynucleotide complementary to it.

Vectors, Host Cells, and Expression Systems

The invention also relates to vectors that comprise a polynucleotide or polynucleotides of the invention, host cells that are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the invention

Recombinant STAAU_R4 and bacteriophage polypeptides of the present invention may be prepared by processes well known to those skilled in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems that comprise a STAAU_R4 or bacteriophage polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems, and to the production of polypeptides of the invention by recombinant techniques.

For recombinant production of a STAAU_R4 polypeptide of the invention, host cells can be genetically engineered to incorporate expression systems or portions thereof or polynucleotides of the invention. Representative examples of appropriate hosts include bacterial cells (Gram positive and Gram

negative), fungal cells, insect cells, animal cells and plant cells. Polynucleotides are introduced to bacteria by standard chemical treatment protocols, such as the induction of competence to take up DNA by treatment with calcium chloride (Sambrook et al., supra). Introduction of polynucleotides into fungal (e.g., yeast) host cells is effected, if desired, by standard chemical methods, such as lithium acetate-mediated transformation.

A great variety of expression systems are useful to produce polypeptides of the invention. Such vectors include among others, chromosomal-, episomal- and virus-derived vectors. For example, vectors derived from bacterial plasmids, from bacteriophages, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses, and from vectors derived from combinations thereof, are useful in the invention.

Polypeptides of the invention are recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid or urea extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Well known techniques for refolding may be employed to regenerate an active conformation when the polypeptide is denatured during isolation and/or purification.

Diagnostic, Prognostic, Serotyping, and Mutation Assays

This invention is also related to the use of STAAU_R4 polynucleotides and polypeptides of the invention for use as diagnostic reagents. Detection of *S. aureus* STAAU_R4 polynucleotides and/or polypeptides in a eukaryote, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of disease, staging of disease or response of an infectious organism to drugs. Eukaryotes, particularly mammals, and especially humans, particularly those infected or suspected to be infected with an organism comprising the *S. aureus* STAAU_R4 gene or protein, may be detected at the

nucleic acid or amino acid level by a variety of well known techniques as well as by methods provided herein.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Polynucleotides from any of these sources, particularly DNA or RNA, may be used directly for detection or may be amplified enzymatically by using PCR or any other amplification technique prior to analysis. RNA, particularly mRNA, cDNA and genomic DNA may also be used in the same ways. Using amplification, characterization of the species and strain of infectious or resident organism present in an individual, may be made by an analysis of the genotype of a selected polynucleotide of the organism. Deletions and insertions can be detected by a change in size of the amplified product in comparison to a genotype of a reference sequence selected from a related organism, preferably a different species of the same genus or a different strain of the same species.

Point mutations can be identified by hybridizing amplified DNA to labeled STAAU_R4 polynucleotide sequences. Perfectly or significantly matched sequences can be distinguished from imperfectly or more significantly mismatched duplexes by DNase or RNase digestion, for DNA or RNA respectively, or by detecting differences in melting temperatures or renaturation kinetics. Polynucleotide sequence differences may also be detected by alterations in the electrophoretic mobility of polynucleotide fragments in gels as compared to a reference sequence. This may be carried out with or without denaturing agents. Polynucleotide differences may also be detected by direct DNA or RNA sequencing. See, for example, Myers *et al*, (1985) Science 230, 1242. Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase, V1 and S1 protection assay or a chemical cleavage method. See, for example, Cotton *et al.*, (1985) Proc. Natl. Acad. Sci., USA 85, 4397-4401.

In another embodiment, an array of oligonucleotide probes comprising STAAU_R4 nucleotide sequence or fragments thereof can be

constructed to conduct efficient screening of, for example, genetic mutations, serotype, taxonomic classification or identification. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability [see, for example, Chee et al., (1996) Science 274, 610].

Thus in another aspect, the present invention relates to a diagnostic kit which comprises: (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1, or a fragment thereof; (b) a nucleotide sequence complementary to that of (a); (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO: 2 or a fragment thereof; or (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO: 2 or fragment thereof.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, among others.

This invention also relates to the use of STAAU_R4 polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of a polynucleotide of the invention, preferably, SEQ ID NO: 1, which is associated with a disease or pathogenicity will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, a prognosis of a course of disease, a determination of a stage of disease, or a susceptibility to a disease, which results from under-expression, over-expression or altered expression of the polynucleotide. Organisms, particularly infectious organisms, carrying mutations in such polynucleotide may be detected at the polynucleotide level by a variety of techniques, such as those described elsewhere herein.

The STAAU_R4 nucleotide sequences of the present invention are also valuable for organism chromosome identification. The sequence is specifically targeted to, and can hybridize with, a particular location on an organism's chromosome, particularly to a *S. aureus* chromosome. The mapping of relevant sequences to chromosomes according to the present

invention may be an important step in correlating those sequences with pathogenic potential and/or an ecological niche of an organism and/or drug resistance of an organism, as well as the essentiality of the gene to the organism. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data may be found on-line in a sequence database. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through known genetic methods, for example, through linkage analysis (coinheritance of physically adjacent genes) or mating studies, such as by conjugation.

The differences in a polynucleotide and/or polypeptide sequence between organisms possessing a first phenotype and organisms possessing a different, second different phenotype can also be determined. If a mutation is observed in some or all organisms possessing the first phenotype but not in any organisms possessing the second phenotype, then the mutation is likely to be the causative agent of the first phenotype.

Polypeptides and polynucleotides for prognosis, diagnosis or other analysis may be obtained from a putatively infected and/or infected individual's bodily materials. Particularly DNA or polynucleotides, from any of these sources may be used directly for detection or may be amplified enzymatically using PCR or other amplification technique with oligonucleotide amplification primers derived from the polynucleotide sequence of *S. aureus* STAAU_R4. RNA, particularly mRNA, or RNA reverse transcribed to cDNA, is also useful for diagnostics. Following amplification of a *S. aureus* STAAU_R4-related polynucleotide from a sample, characterization of the species and strain of infecting or resident organism is made by an analysis of the amplified polynucleotide relative to one or more reference polynucleotides or sequences relative to a standard from a related organism (i.e. a known strain of *S. aureus*).

The invention further provides a process for diagnosing bacterial infections such as those caused by *S. aureus*, the process comprising

determining from a sample derived from an individual, such as a bodily material, an increased level of expression of a polynucleotide having a sequence disclosed in SEQ ID NO: 1 relative to a sample taken from a non-diseased individual. Increased or decreased expression of a STAAU_R4 polynucleotide can be measured using any one of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods, and spectrometry.

In addition, a diagnostic assay in accordance with the invention for detecting over-expression of STAAU_R4 polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example. Assay techniques that can be used to determine levels of a *S. aureus* STAAU_R4 polypeptide, in a sample derived from a host, such as a bodily material, are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis, antibody sandwich assays, antibody detection and ELISA assays.

Gridding and Polynucleotide Subtraction of *S. aureus* Genomic Sequences

The STAAU_R4 polynucleotides of the invention may be used as components of polynucleotide arrays, preferably high density arrays or grids. These high density arrays are particularly useful for diagnostic and prognostic purposes. For example, a set of spots each comprising a different gene, and further comprising a polynucleotide or polynucleotides of the invention, may be used for probing, such as hybridization or nucleic acid amplification, using a probe obtained or derived from a bodily sample, to determine the presence a particular polynucleotide sequence or related sequence in an individual.

Antibodies Specific for *S. aureus* Peptides or Polypeptides

The STAAU_R4 polypeptides and polynucleotides of the invention or variants thereof, or cells expressing them are useful as immunogens to produce antibodies immunospecific for such polypeptides or polynucleotides, respectively.

In certain preferred embodiments of the invention there are provided antibodies against *S. aureus* STAAU_R4 polypeptides or polynucleotides encoding them. Antibodies against STAAU_R4-polypeptide or STAAU_R4-polynucleotide are useful for treatment of infections, particularly bacterial infections.

Antibodies generated against the polypeptides or polynucleotides of the invention are obtained by administering the polypeptides and/or polynucleotides of the invention or epitope-bearing fragments of either or both, analogues of either or both, or cells expressing either or both, to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique known in the art that provides antibodies produced by continuous cell line cultures is useful. Examples include various techniques, such as those in Kohler, G. and Milstein, C., Nature 256: 495-497 (1975); Kozbor *et al.*, Immunology Today 4: 72 (1983); and Cole *et al.*, pg. 96-96 in Monoclonal Anbitodies and Cancer Therapy, Alan R. Liss, Inc. (1985).

Techniques for the production of single chain antibodies (US Patent No: 4,946,968) can be adapted to produce single chain antibodies to polypeptides or polynucleotides of this invention. Also, transgenic mice, or other mammals, are useful to express humanized antibodies immunospecific to the polypeptides or polynucleotides of the invention.

When antibodies are administered therapeutically, the antibody or variant thereof is preferably modified to make it less immunogenic in the individual. For example, if the individual is human the antibody is most preferably "humanized," where the complementarity determining region or regions of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones *et al.* (1986), Nature 321, 522-525 or Tempest *et al.*, (1991) Biotechnology 9, 266-273.

Alternatively, phage display technology is useful to select antibody genes with binding activities towards a STAAU_R4 polypeptide of the invention. In one possible scheme, antibody fragments specific for *S. aureus*

STAAU_R4 are selected from an immune library of antibody genes expressed as fusions with coat protein of filamentous phage. Alternatively, naive libraries are screened by phage display techniques to identify genes encoding antibodies specified for STAAU_R4 or from naive libraries [McCafferty, *et al.*, (1990), Nature 348, 552-554; Marks, *et al.*, (1992) Biotechnology 10, 969-783; a recent reference is de Haard *et al.* (1999) J. Biol. Chem. 274: 18218-18230]. The ability to recover, for various targets, antibodies with subnanomolar affinities obviates the need for immunization. The affinity of these antibodies can also be improved by, for example, chain shuffling [Clackson *et al.*, (1991) Nature 352: 628].

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptides or polynucleotides of the invention, for example to purify the polypeptides or polynucleotides by immunoaffinity chromatography.

A variant polypeptide or polynucleotide of the invention, such as an antigenically or immunologically equivalent derivative or a fusion protein of the polypeptide is also useful as an antigen to immunize a mouse or other animal such as a rat or chicken. A fused protein provides stability to the polypeptide acting as a carrier, or acts as an adjuvant or both. Alternatively, the antigen is associated, for example by conjugation, with an immunogenic carrier protein, such as bovine serum albumin, keyhole limpet haemocyanin or tetanus toxoid. Alternatively, when antibodies are to be administered therapeutically, alternatively a multiple antigenic polypeptide comprising multiple copies of the polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

In accordance with an aspect of the invention, there is provided the use of a STAAU_R4 polynucleotide of the invention for therapeutic or prophylactic purposes, in particular genetic immunization. The use of a STAAU_R4 polynucleotide of the invention in genetic immunization preferably employs a suitable delivery method such as direct injection of plasmid DNA into

muscles [Wolff *et al.*, Hum Mol Genet (1992) 1: 363, Manthorpe *et al.*, Hum. Gene Ther. (1983) 4: 419], delivery of DNA complexed with specific protein carriers [Wu *et al.*, JBiol Chem. (1989) 264: 16985], coprecipitation of DNA with calcium phosphate [Benvenisty and Reshef, Proc. Natl. Acad. Sci. USA, (1986) 83: 9551], encapsulation of DNA in various forms of liposomes [Kaneda *et al.*, Science (1989) 243: 375], particle bombardment [Tang *et al.*, Nature (1992) 356:152, Eisenbraun *et al.*, DNA Cell Biol (1993) 12: 791] or *in vivo* infection using cloned retroviral vectors [Seeger *et al.*, Proc. Natl, Acad. Sci. USA (1984) 81: 5849].

Antagonists and Agonists: Assays and Molecules

The invention is based in part on the discovery that STAAU_R4 is a target for the bacteriophage 3A ORF 33 inhibitory factor. Applicants have recognized the utility of the interaction in the development of antibacterial agents. Specifically, the inventors have recognized that 1) STAAU_R4 is a critical target for bacterial inhibition; 2) 3A ORF 33 or derivatives or functional mimetics thereof are useful for inhibiting bacterial growth; and 3) the interaction between STAAU_R4 or fragment thereof of *S. aureus* and 3A ORF 33 may be used as a target for the screening and rational design of drugs or antibacterial agents. In addition to methods of directly inhibiting STAAU_R4 activity, methods of inhibiting STAAU_R4 expression are also attractive for antibacterial activity.

In several embodiments of the invention, there are provided methods for identifying compounds which bind to or otherwise interact with and inhibit or activate an activity or expression of a polypeptide and/or polynucleotide of the invention comprising: contacting a polypeptide and/or polynucleotide of the invention with a compound to be screened under conditions to permit binding to or other interaction between the compound and the polypeptide and/or polynucleotide to assess the binding to or other interaction with the compound, such binding or interaction preferably being associated with a second component capable of providing a detectable signal in response to the binding or interaction of the polypeptide and/or polynucleotide with the compound; and determining

whether the compound binds to or otherwise interacts with and activates or inhibits an activity or expression of the polypeptide and/or polynucleotide by detecting the presence or absence of a signal generated from the binding or interaction of the compound with the polypeptide and/or polynucleotide.

Potential antagonists include, among others, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention and thereby inhibit or extinguish its activity or expression. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing STAAU_R4-induced activities, thereby preventing the action or expression of *S. aureus* STAAU_R4 polypeptides and/or polynucleotides by excluding *S. aureus* STAAU_R4 polypeptides and/or polynucleotides from binding.

Potential antagonists also include a small molecule that binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules [see Okano, (1991) J. Neurochem. 56, 560; see also Oligodeoxynucleotides As Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988), for a description of these molecules]. Preferred potential antagonists include compounds related to and variants of 3A ORF 33 and of STAAU_R4. Other examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, e.g., Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991). Peptide modulators can also be selected by screening large random libraries of all possible peptides of a certain length.

Compounds derived from the polypeptide sequence of 3A ORF 33 could represent fragments representing small overlapping peptides spanning the entire amino acid sequence of these ORFs. Fragments of 3A ORF 33 can be produced as described above.

Certain of the polypeptides of the invention are biomimetics, functional mimetics of the natural *S. aureus* STAAU_R4 polypeptide. These functional mimetics are useful for, among other things, antagonizing the activity of *S. aureus* STAAU_R4 polypeptide or as an antigen or immunogen in a manner described above. Functional mimetics of the polypeptides of the invention include but are not limited to truncated polypeptides. Polynucleotides encoding each of these functional mimetics may be used as expression cassettes to express each mimetic polypeptide. It is preferred that these cassettes comprise 5' and 3' restriction sites to allow for a convenient means to ligate the cassettes together when desired. It is further preferred that these cassettes comprise gene expression signals known in the art or described elsewhere herein.

Screening Assays According to the Invention

It is desirable to devise screening methods to identify compounds which stimulate or which inhibit the function of the STAAU_R4 polypeptide or polynucleotide of the invention. Accordingly, the present invention provides for a method of screening compounds to identify those that modulate the function of a polypeptide or polynucleotide of the invention. In general, antagonists may be employed for therapeutic and prophylactic purposes. It is contemplated that an agonist of STAAU_R4 may be useful, for example, to

enhance the growth rate of bacteria in a sample being cultured for diagnostic or other purposes.

It has been determined that STAAU_R4 is a target for bacteriophage 3A ORF 33 product, which acts as an inhibitory factor. Applicants have recognized the utility of the interaction in the development of antibacterial agents. Polypeptide and/or polynucleotide targets such as STAAU_R4 are critical targets for bacterial inhibition. *S. aureus* bacteriophage 3A ORF 33 or derivatives or functional mimetics thereof are useful for inhibiting bacterial growth and the interaction, binding, inhibition and/or activation which occurs between polypeptides, such as for example STAAU_R4 of *S. aureus* and 3A ORF 33 may be used for the screening and rational design of drugs or antibacterial agents. In addition to methods for directly inhibiting a target such as STAAU_R4 activity, methods of inhibiting a target such as STAAU_R4 expression are also attractive for antibacterial activity.

In preferred embodiments, the method involves the interaction of an inhibitory ORF product or fragment thereof with the corresponding bacterial target or fragment thereof that maintains the interaction with the ORF product or fragment. Interference with the interaction between the components can be monitored, and such interference is indicative of compounds that may inhibit, activate, or enhance the activity of the target molecule.

a. Binding Assays

There are a number of methods of examining binding of a candidate compound to a protein target such as STAAU_R4 and a polypeptide comprising amino acid sequence of SEQ ID NO: 2, or fragment thereof. Screening methods that measure the binding of a candidate compound to a STAAU_R4 polypeptide or polynucleotide, or to cells or supports bearing the polypeptide or a fusion protein comprising the polypeptide, by means of a label directly or indirectly associated with the candidate compound, are useful in the invention.

The screening method may involve competition for binding of a labeled competitor such as 3A ORF 33 or a fragment that is competent to bind STAAU_R4 or fragment thereof.

Non-limiting examples of screening assays in accordance with the present invention include the following [Also reviewed in Sittampalam *et al.* 1997 Curr. Opin. Chem. Biol. 3:384-91]:

i.) Time-Resolved Fluorescence Resonance Energy Transfer (TR-FRET)

A method of measuring inhibition of binding of two proteins using fluorescence resonance energy transfer [FRET; de Angelis, 1999, Physiological Genomics]. FRET is a quantum mechanical phenomenon that occurs between a fluorescence donor (D) and a fluorescence acceptor (A) in close proximity (usually < 100 A of separation.) if the emission spectrum of D overlaps with the excitation spectrum of A. Variants of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* are fused to a polypeptide or protein and serve as D-A pairs in a FRET scheme to measure protein-protein interaction. Cyan (CFP: D) and yellow (YFP: A) fluorescence proteins are linked with STAAU_R4 polypeptide, or a fragment thereof and a 3A ORF 33 polypeptide respectively. Under optimal proximity, interaction between the STAAU_R4 polypeptide and a 3A ORF 33 polypeptide causes a decrease in intensity of CFP fluorescence concomitant with an increase in YFP fluorescence.

The addition of a candidate modulator to the mixture of appropriately labeled STAAU_R4 and 3A ORF 33 polypeptide, will result in an inhibition of energy transfer evidenced by, for example, a decease in YFP fluorescence at a given concentration of 3A ORF 33 relative to a sample without the candidate inhibitor.

An extension of the FRET technology, termed time-resolved FRET (TR-FRET or HTRF [homogeneous time-resolved energy transfer]) lends itself particularly well to identification of protein-protein interactions in the context of high-throughput screening. In brief, TR-FRET constitutes a homogeneous assay method based on the long-lived fluorescence of rare earth cryptates such

as Europium (Eu) and amplification by nonradiative energy transfer to a suitable acceptor such as allophycocyanin (APC). The TR-FRET principle allows double discrimination of the emitted signal through temporal and spectral selectivity. Since the lifetime of fluorescence emission from APC (acceptor) contains a contribution equal to the Eu (donor) lifetime in the presence of nonradiative energy transfer, a long-lived APC acceptor signal can be resolved from its natural prompt fluorescence in the absence of energy transfer. Eu and APC are brought into proximity via a pair of interacting molecules such as polypeptides. To demonstrate interaction between the STAAU_R4 polypeptide, or a fragment thereof, and a 3A ORF 33 polypeptide, the respective polypeptide is labeled by recombinant DNA methodology to contain an N- or C-terminal tag that is recognized by a binding molecule which itself is conjugated to either Eu or APC. A variety of binding molecules may be employed, including an antibody (directed against an epitope) or streptavidin (directed against biotin). Alternatively, one or both of the interacting proteins is conjugated directly to either Eu or APC.

In one of several possible assay formats, STAAU_R4, or a fragment thereof is expressed as a fusion with a polyhistidine tag and is recognized by an anti-polyhistidine Eu antibody conjugate; 3A ORF 33 is expressed as a fusion with GST and is detected by an anti-GST APC antibody conjugate. Under optimal proximity and in the presence of the anti-polyhistidine and anti-GST antibody conjugates, interaction between STAAU_R4, or a fragment thereof, and 3A ORF 33 induces nonradiative, time-resolved energy transfer from Eu to APC, detected optimally at 665 nm.

The addition of a candidate modulator to the mixture of appropriately labeled STAAU_R4 and 3A ORF 33 polypeptide, will result in an inhibition of energy transfer evidenced by, for example, a decease in APC fluorescence at a given concentration of 3A ORF 33 relative to a sample without the candidate inhibitor.

ii.) Fluorescence polarization

Fluorescence polarization measurement is another useful method to quantitate protein-protein binding. The fluorescence polarization value for a fluorescently-tagged molecule depends on the rotational correlation time or tumbling rate. Protein complexes, such as those formed by a *S. aureus* STAAU_R4 polypeptide, or a fragment thereof associating with a fluorescently labeled polypeptide (e.g., 3A ORF 33 or a binding fragment thereof), have higher polarization values than does the fluorescently labeled polypeptide. Inclusion of a candidate inhibitor of the STAAU_R4 interaction results in a decrease in fluorescence polarization relative to a mixture without the candidate inhibitor if the candidate inhibitor disrupts or inhibits the interaction of STAAU_R4 with its polypeptide binding partner. It is preferred that this method be used to characterize small molecules that disrupt the formation of polypeptide or protein complexes.

iii.) Surface plasmon resonance

Another powerful assay to screen for inhibitors of a protein: protein interaction is surface plasmon resonance. Surface plasmon resonance is a quantitative method that measures binding between two (or more) molecules by the change in mass near a sensor surface caused by the binding of one protein or other biomolecule from the aqueous phase (analyte) to a second protein or biomolecule immobilized on the sensor(ligand). This change in mass is measured as resonance units versus time after injection or removal of the second protein or biomolecule (analyte) and is measured using a Biacore Biosensor (Biacore AB) or similar device. STAAU_R4, or a polypeptide comprising a fragment of STAAU_R4 , could be immobilized as a ligand on a sensor chip (for example, research grade CM5 chip; Biacore AB) using a covalent linkage method (e.g. amine coupling in 10 mM sodium acetate [pH 4.5]). A blank surface is prepared by activating and inactivating a sensor chip without protein immobilization. Alternatively, a ligand surface can be prepared by noncovalent capture of ligand on the surface of the sensor chip by means of a peptide affinity

tag, an antibody, or biotinylation. The binding of 3A ORF 33 to STAAU_R4, or a fragment thereof, is measured by injecting purified 3A ORF 33 over the ligand chip surface. Measurements are performed at any desired temperature between 4°C and 37°C. Conditions used for the assay (i.e., those permitting binding) are as follows: 25 mM HEPES-KOH (pH 7.6), 150 mM sodium chloride, 15% glycerol, 1 mM dithiothreitol, and 0.001% Tween 20 with a flow rate of 10 ul/min. Preincubation of the sensor chip with candidate inhibitors will predictably decrease the interaction between 3A ORF 33 and STAAU_R4. A decrease in 3A ORF 33 binding, detected as a reduced response on sensorgrams and measured in resonance units, is indicative of competitive binding by the candidate compound.

iv.) Scintillation Proximity Assay

A scintillation proximity assay (SPA) may be used to characterize the interaction between a *S. aureus* STAAU_R4 polypeptide, or a fragment thereof, for example comprising the amino acid sequence of SEQ ID NO: 2 and another polypeptide. The SPA relies in a solid-phase substrate, such as beads or the plastic of a microtitre plate, into which a scintillant has been incorporated. For the assay, the target protein, for example a *S. aureus* STAAU_R4 polypeptide, is coupled to the beads or to the surface of the plate, either covalently through activated surface chemistries or non-covalently through a peptide affinity tag, an antibody, or biotinylation. Addition of a radiolabeled binding polypeptide, for example [^{32}P]-radiolabeled 3A ORF 33 , results in close proximity of the radioactive source molecule to the scintillant. As a consequence, the radioactive decay excites the scintillant contained within the bead or within the plastic of the plate and detectable light is emitted. Compounds that prevent the association between immobilized *S. aureus* STAAU_R4 polypeptide and radiolabeled 3A ORF 33 will diminish the scintillation signal. The SPA thus represents an example of an ideal technology with which to screen for inhibitors of the STAAU_R4-3A ORF 33 interactions because it is readily adapted to high-

throughput, automated format and because of its sensitivity for detection of protein-protein interactions with $K_D$values in the micromolar to nanomolar ranges.

v.) Bio Sensor Assay

ICS biosensors have been described by AMBRI (Australian Membrane Biotechnology Research Institute; http//www.ambri.com.au/). In this technology, the self-association of macromolecules such as STAAU_R4, or fragment thereof, and bacteriophage 3A ORF 33 or fragment thereof, is coupled to the closing of gramacidin-facilitated ion channels in suspended membrane bilayers and hence to a measurable change in the admittance (similar to impedence) of the biosensor. This approach is linear over six order of magnitude of admittance change and is ideally suited for large scale, high through-put screening of small molecule combinatorial libraries.

vi.) Phage display

Phage display is a powerful assay to measure protein:protein interaction. In this scheme, proteins or peptides are expressed as fusions with coat proteins or tail proteins of filamentous bacteriophage. A comprehensive monograph on this subject is *Phage Display of Peptides and Proteins. A Laboratory Manual* edited by Kay *et al.* (1996) Academic Press. For phages in the Ff family that include M13 and fd, gene III protein and gene VIII protein are the most commonly-used partners for fusion with foreign protein or peptides. Phagemids are vectors containing origins of replication both for plasmids and for bacteriophage. Phagemids encoding fusions to the gene III or gene VIII can be rescued from their bacterial hosts with helper phage, resulting in the display of the foreign sequences on the coat or at the tip of the recombinant phage.

In one example of a simple assay, purified recombinant STAAU_R4 protein, or fragment thereof, could be immobilized in the wells of a microtitre plate and incubated with phages displaying a 3A ORF 33 sequence in fusion with the gene III protein. Washing steps are performed to remove unbound phages and bound phages are detected with monoclonal antibodies directed against phage coat protein (gene VIII protein). An enzyme-linked secondary

antibody allows quantitative detection of bound fusion protein by fluorescence, chemiluminescence, or colourimetric conversion. Screening for inhibitors is performed by the incubation of the compound with the immobilized target before the addition of phages. The presence of an inhibitor will specifically reduce the signal in a dose-dependent manner relative to controls without inhibitor.

It is important to note that in assays of protein-protein interaction, it is possible that a modulator of the interaction need not necessarily interact directly with the domain(s) of the proteins that physically interact. It is also possible that a modulator will interact at a location removed from the site of protein-protein interaction and cause, for example, a conformational change in the STAAU_R4 polypeptide. Modulators (inhibitors or agonists) that act in this manner can be termed allosteric effectors and are of interest since the change they induce may modify the activity of the STAAU_R4 polypeptide.

Testing for inhibitors is performed by the incubation of the compound with the reaction mixtures. The presence of an inhibitor will specifically reduce the signal in a dose-dependent manner relative to controls without inhibitor. Compounds selected for their ability to inhibit interactions between STAAU_R4-3A ORF 33 are further tested in secondary screening assays.

b. Bacterial growth inhibition

Compounds selected for their ability to inhibit interactions between a STAAU_R4 polypeptide and 3A ORF 33 polypeptide or to inhibit the STAAU_R4 activity can be further tested in functional assays of bacterial growth. Cultures of *S. aureus* are grown in the presence of varying concentrations of a candidate compound added directly to the medium or using a vehicle which is appropriate for the delivery of the compound into the cell. For compounds that correspond to polypeptides, the nucleotide sequence encoding said polypeptides can be cloned into a *S. aureus* expression vector containing an inducible promotor. The expression of the polypeptide could be induced following transfection of cells. For example, the polypeptide may include, but is not limited to different 3A ORF 33 -derived fragments.

Following the induction of expression or the addition of compound, the cultures are then incubated for an additional 4 h at 37°C. During that period of time, the effect of inhibitors on bacterial cell growth may be monitored at 40 min intervals, by measuring, for example, the OD_{565} and the number of colony forming units (CFU) in the cultures. The number of CFU is evaluated as follows: cultures are serially diluted and aliquots from the different cultures are plated out on agar plates. Following incubation overnight at 37°C, the number of colonies is counted. Non-treated cultures of *S. aureus* are included as negative control.

In another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for a polypeptide and/or polynucleotide of the present invention; or compounds which decrease or enhance the production of such polypeptides and/or polynucleotides, which comprises: (a) a polypeptide and/or a polynucleotide of the present invention; (b) a recombinant cell expressing a polypeptide and/or polynucleotide of the present invention; (c) a cell membrane associated with a polypeptide and/or polynucleotide of the present invention; or (d) an antibody to a polypeptide and/or polynucleotide of the present invention; which polypeptide is preferably that of SEQ ID NO: 2, and for which the polynucleotide is preferably that of SEQ ID NO: 1.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

It will be readily appreciated by the skilled artisan that a polypeptide and/or polynucleotide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide and/or polynucleotide, by: (a) determining in the first instance the three-dimensional structure of the polypeptide and/or polynucleotide, or complexes thereof; (b) deducing the three-dimensional structure for the likely reactive site(s), binding site(s) or motif(s) of an agonist, antagonist or inhibitor; (c) synthesizing candidate compounds that are predicted to bind to or react with the

deduced binding site(s), reactive site(s), and/or motif(s); and (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors. It will be further appreciated that this will normally be an iterative process, and this iterative process may be performed using automated and computer-controlled steps.

Each of the polynucleotide sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein, upon expression, can be used as a target for the screening of antibacterial drugs. Additionally, the polynucleotide sequences encoding the amino terminal regions of the encoded protein or Shine-Dalgarno or other sequence that facilitate translation of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The invention also provides the use of the polypeptide, polynucleotide, agonist or antagonist of the invention to interfere with the initial physical interaction between a pathogen or pathogens and a eukaryotic, preferably mammalian, host that is responsible for sequelae of infection. In particular, the molecules of the invention may be used: in the prevention of adhesion of bacteria, in particular Gram positive and/or Gram negative bacteria, to eukaryotic, preferably mammalian, extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; to block bacterial adhesion between eukaryotic, preferably mammalian, extracellular matrix proteins and bacterial STAAU_R4 proteins that mediate tissue damage and/or; to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques. In accordance with yet another aspect of the invention, there are provided STAAU_R4 antagonists, preferably bacteriostatic or bacteriocidal antagonists.

The antagonists of the invention may be employed, for instance, to prevent, inhibit and/or treat diseases.

Compositions, kits and administration

In a further aspect of the invention there are provided compositions comprising a STAAU_R4 polynucleotide and/or a *S. aureus* STAAU_R4 polypeptide for administration to a cell or to a multicellular organism.

The present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide and/or polynucleotide, such as the soluble form of a polypeptide and/or polynucleotide of the present invention, antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1% to about 98% by weight of the formulation; more usually they will constitute up to about 80% by weight of the formulation. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other

compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, and the like.

For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.01 mg/kg to 10 mg/kg, typically around 1 mg/kg. The physician in any event will determine the actual dosage that will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

As used herein, the term "in-dwelling device" refers to surgical implants, prosthetic devices and catheters, i.e., devices that are introduced to the body of an individual and remain in position for an extended time. Such devices include, but are not limited to, artificial joints, heart valves, pacemakers, vascular grafts, vascular catheters, cerebrospinal fluid shunts, urinary catheters, continuous ambulatory peritoneal dialysis (CAPD) catheters.

The composition of the invention may be administered by injection to achieve a systemic effect against relevant bacteria shortly before insertion of an in-dwelling device. Treatment may be continued after surgery during the in-body time of the device. In addition, the composition could also be used to broaden perioperative cover for any surgical technique to prevent bacterial wound infections, especially *S. aureus* wound infections.

Many orthopedic surgeons consider that humans with prosthetic joints should be considered for antibiotic prophylaxis before dental treatment that could produce a bacteremia. Deep infection is a serious complication sometimes leading to loss of the prosthetic joint and is accompanied by significant morbidity and mortality. It may therefore be possible to extend the

use of the active agent as a replacement for prophylactic antibiotics in this situation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

Alternatively, the composition of the invention may be used to bathe an indwelling device immediately before insertion. The active agent will preferably be present at a concentration of 1 mg/ml to 10mg/ml for bathing of wounds or indwelling devices.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.5-5 microgram of antigen /kg, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks. With the indicated dose range, no adverse toxicological effects will be observed with the compounds of the invention that would preclude their administration to suitable individuals.

Sequence Databases, Sequences in a Tangible Medium, and Algorithms

Polynucleotide and polypeptide sequences form a valuable information resource with which to determine their 2- and 3-dimensional structures as well as to identify further sequences of homology. These approaches are most easily facilitated by storing the sequence in a computer readable medium and then using the stored data in a known macromolecular structure program or to search a sequence database using well-known searching tools, such as GCC.

The polynucleotide and polypeptide sequences of the invention are particularly useful as components in databases useful for search analyses as well as in sequence analysis algorithms. As used in this section entitled "Sequence Databases, Sequences in a Tangible Medium, and Algorithms," the terms "polynucleotide of the invention" and "polynucleotide

sequence of the invention" mean any detectable chemical or physical characteristic of a polynucleotide of the invention that is or may be reduced to or stored in a tangible medium, preferably a computer readable form. For example, chromatographic scan data or peak data, photographic data or scan data therefrom, called bases, and mass spectrographic data. As used in this section entitled Databases and Algorithms, the terms "polypeptide of the invention" and "polypeptide sequence of the invention" mean any detectable chemical or physical characteristic of a polypeptide of the invention that is or may be reduced to or stored in a tangible medium, preferably a computer readable form. For example, chromatographic scan data or peak data, photographic data or scan data therefrom, and mass spectrographic data.

The invention provides a computer readable medium having stored thereon polypeptide sequences of the invention and/or polynucleotide sequences of the invention. The computer readable medium can be any composition of matter used to store information or data, including, for example, commercially available floppy disks, tapes, chips, hard drives, compact disks, and video disks.

In a preferred embodiment of the invention there is provided a computer readable medium having stored thereon a member selected from the group consisting of: a polynucleotide comprising the sequence of SEQ ID NO:1; a polypeptide comprising the sequence of SEQ ID NO:2; a set of polynucleotide sequences wherein at least one of said sequences comprises the sequence of SEQ ID NO:1: a set of polypeptide sequences wherein at least one of said sequences comprises the sequence of SEQ ID NO:2; a data set representing a polynucleotide sequence comprising the sequence of SEQ ID NO:1; a data set representing a polynucleotide sequence encoding a polypeptide sequence comprising the sequence of SEQ ID NO:2; a polynucleotide comprising the sequence of SEQ ID NO:1; a polypeptide comprising the sequence of SEQ ID NO:2; a set of polynucleotide sequences wherein at least one of said sequences comprises the sequence of SEQ ID NO: 1; a set of polypeptide sequences

wherein at least one of said sequences comprises the sequence of SEQ ID NO: 2; a data set representing a polynucleotide sequence comprising the sequence of SEQ ID NO: 1; a data set representing a polynucleotide sequence encoding a polypeptide sequence comprising the sequence of SEQ ID NO: 2.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

The present invention is illustrated in further detail by the following non-limiting examples.

EXAMPLE 1

Identification of the inhibitory ORF from Staphylococcus aureus bacteriophage 3A

The *S. aureus* propagating strain 3A (PS 3A obtained from the Laboratory Center for Disease Control (CDC) Health Canada, Ottawa, Ontario) was used as a host to propagate phages 3A (also obtained from CDC). The phage was propagated using the agar layer method described by Swanstörm and Adams [Swanström et al. (1951) Proc. Soc. Exptl. Biol. & Med. 78: 372-375]. Phage DNA was prepared from the purified phages as described in Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

3A ORF 33 (Fig. 2; SEQ ID NO: 4) was amplified by polymerase chain reaction (PCR) from their respective phage genomic DNA. The PCR products were gel purified and digested with BamHI-HindIII. Digested PCR products was then ligated into BamHI-HindIII- digested pTM vector, a *S. aureus* vector containing a kanamycin resistance selective marker (Fig. 3A), and used to

transform *S. aureus* strain RN4220 [Kreiswirth *et al.* (1983) Nature 305: 709-712]. In the resulting vectors, pTM 3AORF33 (Fig.3A), expression of the phage ORF is under the control of an arsenite-inducible promoter. Selection of recombinant clones was performed on Luria-Bertani (LB) agar plates containing 30 μg/ml of kanamycin.

Sodium arsenite ($NaAsO_2$) was used to induce gene expression from the *ars* promoter/operator. The effect of expression of phage ORF on bacterial cell growth was then evaluated in functional assays on solid medium and in liquid medium. As shown in Fig. 3B, the induction of expression of phage 3A ORF 33, by plating three independent transformants (clones 1 to 3) on semi-solid medium containing 5 uM sodium arsenite, results in the inhibition of bacterial growth on solid medium compared to plating in the absence of inducer or plating of a control non-inhibitory ORF (phage 44AHJD ORF 114) transformant. As shown in Fig. 3C and D, the density of the liquid culture, as assessed by colony forming units (CFU), for *S. aureus* clones harboring the 3A ORF 33 increased over time under non-induced conditions. Similar growth rates were also observed with transformants harboring a non-inhibitory ORF (labeled as 'non killer' on the graphs) under both induced and non-induced conditions. Each timepoint in both graphs represents the average obtained from three independent transformants of *S. aureus*.

The expression of 3A ORF 33 inhibits the bacterial growth as observed by the reduction in CFU with time for induced cultures. At 4 h following induction, the expression of 3A ORF 33 is cytocidal resulting in a 2.5 log inhibition reduction in the number of CFU compared to the number of CFU initially present in the same culture (Fig. 3C). When colony plating was done in the absence of kanamycin, the antibiotic necessary to maintain the selective pressure for the plasmid encoding 3A ORF 33 (Fig. 3D), the extent of growth inhibition was reduced, resulting in a 1 log cytocidal effect.

EXAMPLE 2

Identification of a *S. aureus* protein targeted by bacteriophage 3A ORF 33

To identify the *S. aureus* protein(s) that interacts with inhibitory ORF 33 of *S. aureus* bacteriophage 3A, GST-fusion or polyhistidine-fusion of 3A ORF 33 were generated. The recombinant protein was purified and utilized to make a GST/3A ORF 33 affinity column. Cellular extracts prepared from *S. aureus* cells were incubated with the affinity matrix and the matrix was washed with buffers containing increasing concentrations of salt and different detergents. The protein elution profile was assessed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). A protein of molecular mass ~ 40 kDa, identified as PT40, was specifically eluted from the affinity matrix (Fig. 4 and 5) and was not detected in eluates from the GST negative control column (Fig. 6). Eluted proteins were further characterized to determine the identity of the interacting protein and to validate the interaction of the protein with 3A ORF 33 as described in detail below.

A. Generation of fusion 3A ORF 33 recombinant protein.

ORF 33 from bacteriophage 3A was sub-cloned into pGEX 4T-1 (Pharmacia), an expression vector for in-frame translational fusions with GST and into pET15b MCS2, an expression vector in frame with a polyhistidine (histag). Plasmid pTM 3AORF33 (Fig 3A) was purified on a Qiagen column and digested with *Hin*dIII, treated with Klenow fragment of *E. coli* DNA polymerase, and digested with *Bam*HI. The DNA restriction product containing 3A ORF 33 was gel purified by QiAquick spin column (Qiagen) and ligated into pGEX 4T-1 and pET15b MCS2 expression vectors (prepared by digestion with *Sal*I, treatment with Klenow fragment of *E. coli* DNA polymerase, and digestion with *Bam*HI, followed by gel purification by QiAquick spin columns). Recombinant expression vectors were identified by restriction enzyme analysis of plasmid minipreps. Large-scale DNA preparations were performed with Qiagen columns, and the resulting plasmid was sequenced. Test expressions in *E. coli* cells (BL21(DE3) Gold)

containing the expression plasmids were performed to identify optimal protein expression conditions. *E. coli* cells containing the expression constructs were grown in 2 L Luria-Bertani broth at 37°C to an OD_{600} of 0.4 to 0.6 (1 cm pathlength) and induced with 1 mM IPTG for the optimized time and at the optimized temperature (30°C for 3hrs).

B. Fusion of GST/3A ORF 33 and MCS2/3A ORF 33 protein purifications.

Cells containing GST/3A ORF 33 fusion protein were suspended in 10 ml GST lysis buffer/liter of cell culture (GST lysis buffer: 20 mM Hepes pH 7.2, 500 mM NaCl, 10 % glycerol, 1 mM DTT, 1mM EDTA, 1mM benzamidine, and 1 PMSF) and lysed by French Pressure cell followed by three bursts of twenty seconds with an ultra-sonicator at 4°C. The lysate was centrifuged at 4°C for 30 minutes at 10 000 rpm in a Sorval SS34 rotor. The supernatant was applied to a 4 ml glutathione sepharose column pre-equilibrated with lysis buffer and allowed to flow by gravity. The column was washed with 10 column volumes of lysis buffer and eluted in 4 ml fractions with GST elution buffer (20 mM Hepes pH 8.0, 500 mM NaCl, 10 % glycerol, 1 mM DTT, 0.1mM EDTA, and 25 mM reduced glutathione). The fractions were analyzed by 15% SDS-PAGE (Laemmli) and visualized by staining with Coomassie Brilliant Blue R250 stain to assess the amount of eluted GST/3A ORF 33 protein.

Cells containing MCS2/3A ORF 33 (histag) fusion proteins were suspended in 10 ml lysis buffer/liter of cell culture with histag lysis buffer (20 mM Hepes pH 7.5, 500 mM NaCl, 10% glycerol, 1mM benzamidine, and 1 mM PMSF) and lysed by passage through a French Pressure cell followed by three bursts of twenty seconds with an ultra-sonicator at 4°C. The lysate was centrifuged at 4°C for 30 minutes at 10 000 rpm in a Sorval SS34 rotor. The supernatant was applied to a 4 ml metal chelate sepharose column pre-equilibrated with Ni^{2+}, and washed with histag lysis buffer and allowed to flow by gravity. The column was washed with 10 column volumes of histag lysis buffer containing 5 mM imidazole, 6 column volumes of histag lysis buffer containing 50 mM imidazole and eluted in 4 ml fractions with histag elution buffer (20 mM Hepes

pH 7.5, 500 mM NaCl, 500 mM imidazole, and 10% glycerol). The fractions were analyzed by 15% SDS-PAGE (Laemmli) and visualized by staining with Coomassie Brilliant Blue R250 stain.

C. *S. aureus* extract preparation.

A *Staphylococcus aureus* extract was prepared from the cell pellets using lysostaphin digestion followed by sonication and nuclease digestion. The cell pellet (2.9g) was suspended in 8 ml of 20 mM Hepes pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 1000 units of lysostaphin, 0.5 mg RNase A, 750 units micrococcal nuclease, and 375 units DNase 1. The cell suspension was incubated at 37°C for 30 minutes, cooled to 4°C, and made up to a final concentration of 1 mM EDTA and 500 mM NaCl. The lysate was sonicated on ice using three bursts of 20 seconds each. The lysate was centrifuged at 20 000 rpm for 1 hr in a Ti70 fixed angle Beckman rotor. The supernatant was removed and dialyzed overnight in a 10 000 Mr dialysis membrane against Affinity Chromatography Buffer (ACB; 20 mM Hepes pH 7.5, 10 % glycerol, 1 mM DTT, and 1 mM EDTA) containing 100 mM NaCl, 1mM benzamidine, and 1 mM PMSF. The dialyzed protein extract was removed from the dialysis tubing and frozen in one ml aliquots at –70 °C.

D. Affinity column preparation.

GST, GST/3A ORF 33, and MCS2/3A ORF 33 were dialyzed overnight against ACB containing 1 M NaCl. Protein concentrations were determined by Bio-Rad Protein Assay and proteins were crosslinked to Affigel 10 resin (Bio-Rad) at protein/resin concentrations of 0, 0.1, 0.5, 1.0, and 2.0 mg/ml. The crosslinked resin was sequentially incubated in the presence of ethanolamine and bovine serum albumin (BSA) prior to column packing and equilibration with ACB containing 100 mM NaCl. *S. aureus* extracts were centrifuged at 4°C in a micro-centrifuge for 15 minutes and diluted to 5 mg/ml with ACB containing 100 mM NaCl. 400 ul of extract was applied to 40 ul columns containing 0, 0.1, 0.5, 1.0, and 2.0 mg/ml ligand and ACB containing 100 mM NaCl (400 ul) was applied

to an additional column containing 2.0 mg/ml ligand. The columns were washed with ACB containing 100 mM NaCl (400 ul) and sequentially eluted with ACB containing 0.1% Triton X-100 and 100 mM NaCl (100 ul), ACB containing 1 M NaCl (160 ul), and 1% SDS (160 ul). 80 ul of each eluate was resolved by 16 cm 14 % SDS-PAGE [Laemmli, U.K. (1970) Nature 227: 680-685] and the protein was visualized by silver stain.

E. Affinity chromatography

One candidate interacting protein of 40 kDa (PT40) was observed in the 1M NaCl eluates of the GST/3A ORF 33 chromatography experiment (Fig. 4). The PT40 protein was also observed in the 1M NaCl eluates of the MCS2/3A ORF 33 experiment (Fig. 5). This candidate protein was not observed in the GST control affinity chromatography experiment. An estimation of the relative abundance of PT40 protein in the *Staphylococcus aureus* extract relies upon the assumption that nearly quantitative recovery of the candidate interacting protein has occurred during the affinity chromatography. Affinity chromatography experiments with the 5 mg/ml extract using ligands MCS2/3A ORF 33 and GST/3A ORF 33 yielded similar quantities of PT40 with approximately 25 ng of PT40 in the 1M NaCl eluate of the 2.0 mg/ml column. Affinity chromatography experiments with the 10 mg/ml extract using ligands MCS2/3A ORF33 and GST/3A ORF 33 yielded similar quantities of PT40 with approximately 70 ng of PT40 in the 1M NaCl eluate of the 2.0 mg/ml column. Although protein quantitation from silver stained SDS-PAGE gels is only approximate, the estimated abundance of PT40 in the extracts is approximately 0.005%.

F. Identification of *S. aureus* STAAU R4 as an 3A ORF 33 interacting protein

The candidate protein PT40 was excised from the SDS-PAGE gels and prepared for tryptic peptide mass determination by MALDI-ToF mass spectrometry. [Qin, J., *et al.* (1997) *Anal. Chem.* 69, 3995-4001]. As exemplified in Fig. 7, high quality mass spectra were obtained. The PT40 proteins observed in the two affinity chromatography experiments (eluates presented in Figs. 4 and

5) were identical as determined by the masses of the tryptic peptides. The PT40 band was identified as an open reading frame found in Contig782 of the University of Oklahoma genome sequencing project database (http://www.genome.ou.edu/staph.html) (herein referred as 'STAAU_R4'). G. Bioinformatic analysis of STAAU_R4

Sequence homology (BLAST) and Hidden Markov Model (HMM) searches were then carried out using an implementation of both programs. Downloaded public databases used for sequence analysis include:

i) non-redundant GenBank (nr) (www.ncbi.nlm.nih.gov)

ii) pdbaa database (www.ncbi.nlm.nih.gov)

iii) PRODOM (http://protein.toulouse.inra.fr/protein.html)

iv) Swissprot and TREMBL (www.expasy.ch)

v) Block plus and Block prints (http://blocks.fhcrc.org)

vi) Pfam (http://pfam.wustl.edu)

vii) Prosite (www.expasy.ch)

As shown in Fig.8, the results of the BLAST searches performed on STAAU_R4 revealed similarity to fatty acid/phospholipid synthesis PlsX protein from a variety of bacteria, including *B. subtilis*. In particular, the result of the global optimal alignment, presented in Fig. 9A, reveals a 53% identity and a 71% similarity between the amino acid sequences of STAAU_R4 and *B. subtilis* PlsX (gi|6686325|sp|P71018|PLSX_BACSU: FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLSX).

A software program was developed and used on the *S. aureus* sequence to identify the start codon of the candidate ORF. The software scans the primary nucleotide sequence for an appropriate start codon. Two possible selections can be made for defining the nature of the start codon; a) selection of ATG or b) selection of ATG or GTG. The analysis also involved the identification of a ribosomal binding site (RBS) sequence that consists on finding the best ungapped alignment of the Shine-dalgarno sequence '***aaaggaggt***' within the 25

nucleotides long sequence located upstream of the start codon. Position dependent scores are as followed:

a a a g g a g g t

2 3 3 6 6 4 6 5 2

As shown in Fig. 9B, the analysis of the *S. aureus* stop codon-to-stop codon DNA region containing the STAAU_R4 revealed the presence of two possible ATG start codons associated with a predicted RBS sequence with a score of 70 % of the perfect match. Based on the sequence alignment of STAAU_R4 with other PlsX polypeptides, exemplified for *B. subtilis* PlsX in Fig. 9A, the start codon of STAAU_R4 is predicted to correspond to the first methionine downstream of the RBS sequence identified in bold on Fig 9B.

EXAMPLE 3

Affinity blotting interaction between STAAU_R4 and 3A ORF 33.

The bacterial STAAU_R4 and 3A ORF 33 interaction was also confirmed by an affinity blotting assay. The probe was made of the STAAU_R4 protein by incubation of the heart muscle kinase (HMK)-tagged STAAU_R4 GST-fusion with [^{32}P]-ATP and HMK. The labeled probe was incubated with the blocked membrane bearing the immobilized inhibitory phage ORF, and the remaining signal after extensive washing is detected by exposure to X-ray film.

A. Generation of GST/3A ORF 33 and GST/STAAU_R4recombinant polypeptides for affinity blotting analysis.

The pGEX-6P1 harboring GST in fusion with PreScission protease cleavage site was purchased from Pharmacia Amersham Biotech. The pGEX-6PK was obtained by cloning synthetic annealed oligonucleotides corresponding to the heart muscle kinase (HMK) phosphorylation site (SEQ ID NO: 5: 5'- GATCTCGTCGTGCATCTGTTGGATCCCCGGAATTCCCGGG-3' And SEQ ID NO: 6: 5'-TCGACCCGGGAATTCCGGGGATCCAACAGAT-GCACGACGA-3') into the unique *Bam*HI-*Sal*I [Kaelin *et al.* (1992) Cell 70: 351-

364]. As a consequence, the HMK site is cloned in fusion with GST and is followed by unique *Bam*HI, *Eco*RI and *Sal*I cloning sites. Insertion of the duplex was confirmed by DNA sequence analysis and the plasmid is referred to as pGEX-6PK (Fig.10C).

As shown in Fig 10B, bacteriophage 3A ORF 33 was sub-cloned into pGEX-6PK by *EcoRI* and *XhoI* digestion from a pGADT7 vector (Clontech Laboratories) harboring the 3A ORF 33 sequence. The DNA fragment containing 3A ORF 33 was gel purified by Qiagen column and ligated in frame at the 3'-end of the HMK labeling site into pGEX–6PK (which had been previously digested with *EcoRI* and *SalI*) to generate pGEX–6PK3A ORF 33. Recombinant expression vectors were identified by restriction enzyme analysis of plasmid minipreps, large-scale DNA preparations were performed with Qiagen columns.

As shown in Fig 10A, full-length STAAU_R4 was amplified from *S. aureus* genomic DNA by PCR using the sense strand primer targeting the initiation codon preceded by a *Bam*HI restriction site (5'-cgggatccATGGTTAAATTAGCAATTGAT-3') (SEQ ID NO: 7) and the antisense oligonucleotide targeting the stop codon preceded by a *Xho*I restriction site (5'-ccgctcgagTTACTCATTTGATTCACCTAC-3') (SEQ ID NO: 8). The initiation codon used to amplified STAAU_R4 (SEQ ID NO: 1) corresponds to the predicted start codon (Fig. 9B) based on the sequence alignment of STAAU_R4 with other PlsX polypeptides, exemplified for *B. subtilis* PlsX in Fig. 9A. The digested PCR product was purified using the Qiagen PCR purification kit, ligated in frame at the 3'-end of the HMK labeling site into *BamH*I and *SalI* digested pGEX-6PK vector and used to transform *E. coli* strain BL21 (Amersham-Pharmacia). The sequence integrity of STAAU_R4 polypeptides fused to GST was verified directly by DNA sequencing.

B. Purification and radiolabeling of GST-fusion proteins

Expression of the GST/3A ORF 33 and GST/STAAU_R4 recombinant proteins from the plasmid pGEX-6P was induced by the addition of 1 mM of isopropyl-1-thio-β-D-galactosidase (IPTG) to a culture at OD_{600} ~0.5. The

protein was expressed at 37°C for 3h and the cells were harvested by centrifugation at 5,000 rpm in a JA-10 rotor (Beckman) for 15 min at 4°C. The bacterial pellet was resuspended in 100 ml of ice-cold phosphate buffer saline (PBS), divided into 4 aliquots and centrifuged as above. Each aliquot was resuspended in 7 ml of STE (10 mM Tris pH 8.0, 1 mM EDTA, 150 mM NaCl and 0.1 mg/ml lysozyme). After incubation for 15 min, 10 mM dithiothreithol (Gibco BRL) and 1.4% Sarkosyl (Sigma) were added and the lysed cells were submitted to three bursts of sonication of 20 seconds on ice.

The cell lysate was centrifuged at 16,000 rpm in a JA-20 rotor (Beckman) for 20 min at 4°C and the supernatant was treated with 2% Triton X-100 (Sigma) in a total volume of 20 ml for 30 min at room temperature with end-over-end rotation. The lysate was centrifuged at 16,000 rpm in an JA-20 rotor for 20 min at 4C° and the supernatant was incubated with 1 ml of glutathione Sepharose-4B beads (Amersham-Pharmacia) for 60 min at 4C°. Beads were washed extensively with PBS, transferred to an eppendorf tube and the pure proteins were cleaved from the GST portion by digestion with 40 Units of PreScission protease (Pharmacia –Amersham) in 500 µl of 50 mM Tris pH 7.0, 150 mM NaCl, 1 mM EDTA and 1 mM DTT. After 5 hrs incubation at 4C° with end-over-end rotation, samples were centrifuged for 5 min at 13,000 g in a microfuge and the supernatant was collected. Protein concentrations were determined using the Biorad reagent according to the manufacturer's instructions (Biorad) and the proteins were stored at –80°C until use. Proteins were analyzed by 12% SDS-PAGE and visualized by Coommassie Brilliant Blue R-250 staining.

For radiolabeling with [^{32}P]-ATP, 4 µg of GST-cleaved STAAU_R4 polypeptide were incubated with 50 Units of catalytic sub-unit of cAMP dependent protein kinase "Heart Muscle Kinase" (Sigma) in a total volume of 100 µl containing 200 mM Tris pH 7.5, 1 M NaCl, 120 mM $MgCl_2$, 10 mM DTT and 50 µCi of [γ^{32}P]-ATP (3000 ci/mmol) (NEN/Mandel) for 30 min at room temperature. To remove free nucleotides, the proteins were applied to Sephadex-G50 NICK columns (Amersham-Pharmacia) and eluted with Z-buffer (25 mM

Hepes pH 7.7, 12.5 mM $MgCl_2$, 20% Glycerol, 100 mM KCl & 1 mM DTT and the incorporation of γ^{32}P-ATP was determined by counting on a liquid scintillation counter.

C. Affinity blotting assay:

Increasing amounts (0.125, 0.250, 0.5 and 1.0 µg respectively lanes 1, 2, 3 and 4 in Fig 10D) of GST-cleaved 3A ORF 33 were resolved on a sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-PAGE) gels and blotted onto nitrocellulose membrane (Millipore). Immobilized proteins were denatured by incubating the membrane with 6M urea in HBB buffer (25 mM Hepes-KOH pH 7.7, 25 mM NaCl, 5 mM $MgCl_2$, 1mM DTT) for 20 min at 4°C. The proteins were renatured *in situ* by a progressive dilution of urea in HBB buffer. The membrane was blocked for at least 1 hr at 4°C with 5% milk in HBB supplemented with 0.05% NP-40 and for 45 min in 1% milk in HBB supplemented with 0.05% NP-40. The hybridization was performed for overnight at 4°C in hybridization buffer (20 mM Hepes-KOH pH 7.7, 75 mK KCl, 0.1 mM EDTA 2.5 mM $MgCl_2$, 0.05% NP-40 and 1% milk) containing 250,000 cpm/ml of [^{32}P]-ATP labelled STAAU_R4 polypeptide as a probe. The membrane was washed 3 times for 10 min with hybridization buffer and exposed to x-ray film (Fig 10D). A band corresponding to the [^{32}P]-ATP labelled STAAU_R4 polypeptide was detected at a position on the blot that corresponds to 3A ORF 33 (18 kDa).

CONCLUSION

By virtue of the interaction between the inhibitory bacteriophage 3A ORF 33 and the STAAU_R4, the STAAU_R4 gene and its gene product have thus been identified as novel bacterial targets for the screening and identification of anti-bacterial agents and more particularly for anti *S. aureus* agents. The present invention also provides novel diagnosis, prognosis and therapeutic methods based on STAAU_R4, and/or bacteriophage 3A ORF 33 and/or a compound identified in accordance with the present invention.

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified without departing from the spirit and nature of the subject invention as defined in the appended claims.

WHAT IS CLAIMED IS:

1. An isolated, purified or enriched polypeptide consisting of the amino acid sequence of SEQ ID NO: 2.

2. An isolated, purified or enriched polypeptide having at least 55% identity to the amino acid sequence of claim 1.

3. An isolated, purified or enriched polypeptide having at least 60% identity to the amino acid sequence of claim 1.

4. An isolated, purified or enriched polypeptide having at least 75% similarity to the amino acid sequence of claim 1.

5. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or biologically active fragment or variant thereof.

6. An isolated, purified or enriched polypeptide comprising the amino acid sequence having at least 55% identity to the amino acid sequence of claim 5.

7. An isolated, purified or enriched polypeptide comprising the amino acid sequence having at least 60% identity to the amino acid sequence of claim 5.

8. An isolated, purified or enriched polypeptide of at least 50 amino acids in length having at least 55 % identity to the amino acid sequence of SEQ ID NO: 2.

9. An isolated, purified or enriched polynucleotide consisting of the nucleotide sequence as set forth in SEQ ID NO: 1, or the complement of said polynucleotide sequence.

10. An isolated, purified or enriched polynucleotide sequence encoding the polypeptide of claim 5 or the complement of said polynucleotide sequence.

11. An isolated, purified or enriched polynucleotide comprising SEQ ID NO: 1 or the complement thereof.

12. An isolated, purified or enriched polynucleotide comprising a nucleotide sequence having at least 65% identity to the sequence as set forth in SEQ ID NO: 1; or the complement of said polynucleotide.

13. An isolated, purified or enriched polynucleotide comprising a nucleotide sequence encoding any one of the polypeptides from claim 1 to 8.

14. A method for identifying a compound active on a biologically active STAAU_R4 polypeptide comprising the amino acid sequence of SEQ ID NO: 2, or a biologically active variant thereof, said method comprising:

contacting said STAAU_R4 polypeptide with a candidate compound, and

detecting one of: a binding of said compound to said STAAU_R4 polypeptide; and a biological activity thereof,

wherein said binding of the compound to said STAAU_R4 polypeptide or a decrease in said biological activity thereof in the presence of said candidate compound relative to same in the absence thereof, is indicative that said candidate compound is a compound that is active on the STAAU_R4 polypeptide.

15. The method of claim 14, wherein said STAAU_R4 polypeptide is contacted simultaneously with said bacteriophage polypeptide sequence.

16. The method of claim 14 or 15, wherein said binding or said decrease in biological activity is performed in the presence and absence of said candidate compound.

17. The method of one of claims 14 to 16, wherein said bacteriophage polypeptide sequence is selected from the group consisting of:

a) SEQ ID NO:4; and

b) a fragment or variant of a) that specifically binds with SEQ ID NO:2.

18. The method of one of claims 14 to 17, wherein said detecting comprises measuring the binding of a candidate compound to said STAAU_R4 polypeptide, wherein the compound is directly or indirectly detectably labeled.

19. A method for identifying a compound active on one of a STAAU_R4 polypeptide and a bacteriophage polypeptide which specifically interacts with same comprising:

contacting said STAAU_R4 polypeptide which comprises the amino acid sequence of SEQ ID NO: 2, or variant thereof, and said bacteriophage polypeptide which is selected from the group consisting of:

a) SEQ ID NO:4; and

b) a fragment or variant of a), wherein said fragment or variant of a) maintains its biological activity;

with a candidate compound; and

detecting a biological activity of said STAAU_R4 polypeptide and/or said bacteriophage polypeptide, wherein a decrease in the biological activity thereof in the presence of the candidate compound is indicative that said candidate compound is a compound that is active on one of said STAAU_R4 and/or bacteriophage polypeptide.

20. The method of claim 19, which identifies a compound active on STAAU_R4.

21. The method of claim 19 or 20, wherein said detecting comprises the act of measuring the binding of said STAAU_R4 polypeptide to said bacteriophage polypeptide wherein said STAAU_R4 polypeptide or said bacteriophage polypeptide is directly or indirectly detectably labeled.

22. A method of identifying a compound that is active on a biologically active STAAU_R4 polypeptide, said method comprising:

contacting a candidate compound with cells expressing said STAAU_R4 polypeptide comprising the amino acids of SEQ ID NO: 2, or a biologically active fragment, and

detecting a STAAU_R4 activity in said cells,

wherein a decrease in said activity in said cells in the presence of said candidate compound is indicative of an inhibition of STAAU_R4 activity by said compound.

23. An agonist or an antagonist of the activity of a STAAU_R4 polypeptide or fragment thereof, or a nucleic acid encoding said polypeptide or fragment thereof, identified with any one of the methods of claims 14 to 22.

24. The method of any one of claims 14-21, wherein said detecting comprises measurement by time-resolved fluorescence resonance energy transfer (TR-FRET).

25. The method of any one of claims 14-21, wherein said detecting comprises measurement of fluorescence polarization changes.

26. The method of any one of claims 14-21, wherein said detecting comprises measurement by surface plasmon resonance.

27. The method of any one of claims 14-21, wherein said detecting comprises a scintillation proximity assay.

28. The method of any one of claims 14-21, wherein said detecting comprises a biosensor assay.

29. The method of any one of claims 14-21, wherein said detecting comprises measurement by phage display.

30. A method of making an antibacterial compound, comprising:

identifying a compound active on a polypeptide comprising the amino acid sequence selected from:

a) SEQ ID NO:2;

b) a biologically active fragment or variant of a)

c) a nucleic acid encoding any of said polypeptide of a) – b); and

synthesizing or purifying said active compound in an amount sufficient to provide a therapeutic effect when administered to an organism infected or at risk of being infected by a bacterium naturally producing said polypeptide, or nucleic acid encoding same.

31. A method for inhibiting a bacterium, comprising contacting the bacterium with a compound active on one of a) a *S. aureus* polypeptide

comprising the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof; and b) a nucleic acid encoding the polypeptide of a).

32. A method for treating or preventing a bacterial infection in an animal suffering from an infection or at risk of suffering therefrom, comprising administering to said animal a therapeutically effective or prophylactically effective amount of a compound active on a *S. aureus* polypeptide comprising the amino acid sequence of one of a) SEQ ID NO: 2, fragment or variant thereof; b) a nucleic acid encoding said amino acid sequence of a).

33. A method of prophylactic treatment to prevent bacterial infection comprising contacting an indwelling device with a compound active on a *S. aureus* polypeptide comprising the amino acid sequence of SEQ ID NO: 2, before its implantation into a mammal, such contacting being sufficient to prevent *S. aureus* infection at the site of implantation.

34. A method of prophylactic treatment to prevent infection of an animal by a bacterium comprising administering to the animal a compound that is active on a *S. aureus* polypeptide comprising one of the amino acid sequence of SEQ ID NO: 2, fragment or variant thereof; or a gene encoding said polypeptide, in an amount sufficient to reduce adhesion of the bacterium to a tissue surface of a tissue of the mammal.

35. The method of any one of claims 14-22, or 24-34, wherein said active compound is selected from the group consisting of a small molecule, a peptidomimetic compound, and a fragment or derivative of a bacteriophage inhibitor protein.

36. The method of any one of claims 14-22, or 24-34, wherein said active compound is a peptide synthesized by an expression system and purified, or is artificially synthesized.

37. The method of claim 36, wherein said compound is selected from the group consisting of:

a) SEQ ID NO:4; and

b) a fragment or variant of a), wherein said fragment or variant thereof maintains its specific binding capability of interacting with one of SEQ ID NO:2, fragment or variant thereof.

38. The method of any one of claims 31, 33, 35, 36 or 37, wherein said contacting is performed *in vitro.*

39. The method of any one of claims 31, 32, 34-36 or 37, wherein said contacting is performed *in vivo* in an animal.

40. The method of any one of claims 31-38, wherein said contacting is performed in combination with existing antimicrobial agents.

41. The method of any one of claims 14-22, 23, 24-29, or 35-37, wherein said STAAU_R4 polypeptide comprises the amino acid sequence as set forth in SEQ ID NO:2, or biologically active fragment or variant thereof.

42. A composition comprising an isolated, purified or enriched bacteriophage 3A ORF 33-encoded polypeptide; and a *S. aureus* STAAU_R4 polypeptide comprising the amino acid sequence seth forth in SEQ ID NO: 2, fragment or variant thereof.

43. A composition comprising a pair of specifically interacting domains, said pair comprising: a STAAU_R4 polypeptide and a polypeptide encoded by a bacteriophage ORF which specifically interacts with said STAAU_R4 polypeptide.

44. The composition of claim 43, wherein said STAAU_R4 polypeptide comprises the amino acid sequence set forth in SEQ ID NO:2, and wherein said bacteriophage ORF comprises the amino acid sequence as set forth in SEQ ID NO:4.

45. The composition of claim 43 or 44, wherein said STAAU_R4 polypeptide comprises the amino acid sequence set forth in SEQ ID NO:2 or biologically active fragment or variant thereof.

46. A process for producing a pharmaceutical composition comprising: a) identifying a compound that is active on a STAAU_R4 polypeptide by performing a screening assay for compounds active on a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or biologically active fragment or variant thereof; and b) mixing the compound identified in a) with a suitable pharmaceutical carrier.

47. The process of claim 46, wherein said bacteriophage ORF comprises the amino acid sequence as set forth in SEQ ID NO:4.

48. The process of claim 46 or 47, wherein said STAAU_R4 polypeptide comprises the amino acid sequence as set forth in SEQ ID NO:2 or biologically active fragment or variant thereof.

49. Use of one of: a) a STAAU_R4 polypeptide comprising the amino acid sequence of SEQ ID NO:2, a biologically active fragment or variant

thereof; b) a composition comprising a pair of specifically interacting domains comprised of a polypeptide of STAAU_R4, biologically active fragment thereof or variant thereof and a polypeptide encoded by a bacteriophage ORF which specifically interacts with STAAU_R4; or c) an assay mixture comprising a first polypeptide which comprises one of i) the amino acid sequence of SEQ ID NO:2, or ii) a biologically active fragment or variant of i); and a second polypeptide encoded by a bacteriophage ORF which specifically interacts with one of i)-ii); for the identification of a compound that is active on a STAAU_R4 polypeptide.

50. A method of diagnosing in an individual an infection with *S. aureus*, comprising:

determining the presence in said individual of a *S. aureus* STAAU_R4 polypeptide.

51. The method of claim 50 wherein said determining step comprises contacting a biological sample of said individual with an antibody specific for an epitope present on a *S. aureus* STAAU_R4 polypeptide.

52. A method of diagnosing in an individual an infection with *Staphylococcus aureus*, comprising:

determining the presence in said individual of a nucleic acid sequence encoding a *S. aureus* STAAU_R4 polypeptide.

53. The method of claim 52 wherein said determining step comprises contacting a nucleic acid sample of said individual with a probe of at least 15 nucleotides in length that hybridizes under stringent hybridization conditions with the sequence of SEQ ID NO:1, or the complement of said probe.

54. An isolated, purified or enriched antibody specific for a polypeptide derived from SEQ ID NO: 2.

SEQ ID NO: 1

>STAAU_R004 DNA sequence 987 nuceotides
ATGGTTAAATTAGCAATTGATATGGGTGGCGACAATGCGCCTGATATCGTATTAGAA
GCCGTACAAAAGGCTGTTGAAGACTTTAAAGATCTAGAAATTATACTTTTCGGTGACGAA
AAAAAGTATAATCTGAACCATGAACGAATCGAATTTAGACATTGTTCTGAAAAGATTGAA
ATGGAAGATGAGCCTGTTAGAGCGATTAAACGTAAAAAAGATAGCTCAATGGTAAAAATG
GCTGAAGCTGTGAAATCTGGTGAAGCAGATGGATGTGTGTCAGCAGGTAATACTGGTGCT
TTAATGTCAGCTGGTTTATTCATTGTTGGACGTATTAAAGGTGTAGCTAGACCGGCTTTA
GTAGTAACATTGCCAACGATTGATGGAAAAGGTTTTGTCTTTTTAGACGTTGGTGCAAAT
GCTGATGCTAAACCTGAACACTTATTACAGTATGCGCAACTAGGGGATATTTATGCTCAA
AAAATTAGAGGTATTGATAATCCGAAAAATCTCATTATTAAATATAGGAACCGAGCCAGCT
AAAGGTAATAGTTTAACGAAAAAATCATATGAGTTATTAAATCATGATCATTCATTGAAT
TTTGTTGGGAATATTGAAGCGAAGACATTAATGGATGGCGATACAGATGTTGTAGTTACC
GATGGCTATACTGGGAACATGGTCCTTAAAAATTTAGAAGGTACTGCAAAATCAATCGGT
AAAATGTTAAAGATACGATTATGAGTAGTACTAAAAATTAATTAGCAGGTGCAATATTG
AAGAAAGATTTAGCTGAATTCGCTAAAAGATGGATTACTCAGAATACGGTGGTTCCGTA
TTATTAGGATTGGAAGGTACTGTAGTTAAAGCACACGGTAGTTCAAATGCTAAAGCTTTT
TATTCTGCAATTAGACAAGCGAAAATCGCAGGAGAACAAAATATTGTACAAACAATGAAA
GAGACTGTAGGTGAATCAAATGAGTAA

FIG. 1

SEQ ID NO: 2

>STAAU_R004 amino acids sequence
MVKLAIDMMGGDNAPDIVLEAVQKAVEDFKDLEIILFGDEKKYNLNHERIEFRHCSEKIE
MEDEPVRAIKRKKDSSMVKMAEAVKSGEADGCVSAGNTGALMSAGLFIVGRIKGVARPAL
VVTLPTIDGKGFVFLDVGANADAKPEHLLQYAQLGDIYAQKIRGIDNPKISLLNIGTEPA
KGNSLTKKSYELLNHDHSLNFVGNIEAKTLMDGDTDVVVTDGYTGNMVLKNLEGTAKSIG
KMLKDTIMSSTKNKLAGAILKKDLAEFAKKMDYSEYGGSVLLGLEGTVVKAHGSSNAKAF
YSAIRQAKIAGEQNIVQTMKETVGESNE

Fig. 1 (cont.)

SEQ ID NO: 3

>3AORF033 nucleotide sequence
ATGGCAATATTAGAAGGTATTTTTTGAAGAATTAAAACTATTAAATAAGAATTTACGTGTG
CTAAATACTGAACTATCAACTGTAGATTCATCAATTGTACAAGAGAAAGTTAAAGAAGCA
CCAATGCCAAAAGATGAAACAGCTCAACTGGAATCAGTTGAAGAAGTTAAGGAAACTTCT
GCTGATTTAACTAAAGATTATGTTTTATCAGTAGGAAAAGAGTTCCTTAAAAAAGCAGAT
ACTTCTGATAAGAAAGAATTTAGAAATAAACTTAACGAACTTGGTGCGGATAAGCTATCT
ACTATCAAAGAAGAGCATTATGAAAAAATTGTTGATTTTATGAATGCGAGAATAAATGCA
TGA

SEQ ID NO: 4

>3AORF033 amino acid sequence
MAILEGIFEELKLLNKNLRVLNTELSTVDSSIVQEKVKEAPMPKDETAQLESVEEVKETS
ADLTKDYVLSVGKEFLKKADTSDKKEFRNKLNELGADKLSTIKEEHYEKIVDFMNARINA

Fig. 2

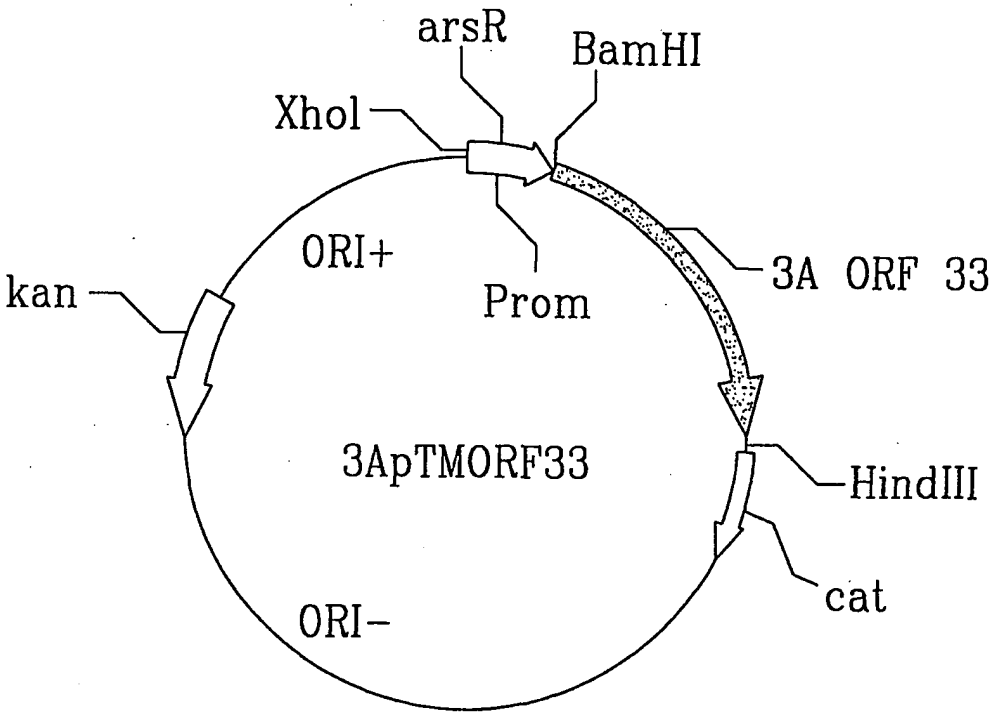

arsR
BamHI
XhoI
ORI+
Prom
3A ORF 33
kan
3ApTMORF33
HindIII
cat
ORI-

FIG. 3A

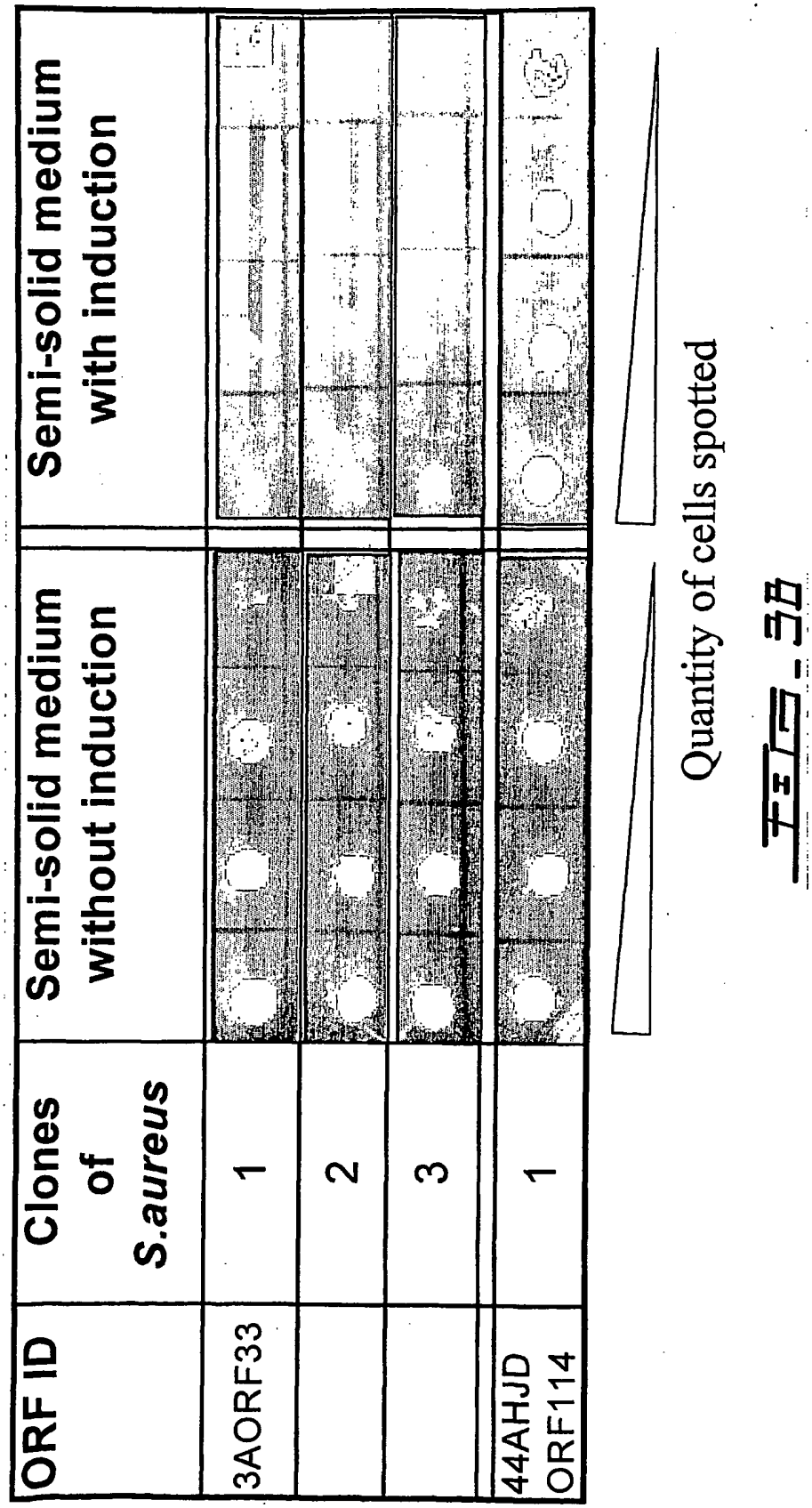

ORF ID
Clones of S.aureus
Semi-solid medium without induction
Semi-solid medium with induction
3AORF33
1
2
3
44AHJD ORF114
1
Quantity of cells spotted

FIG - 3B

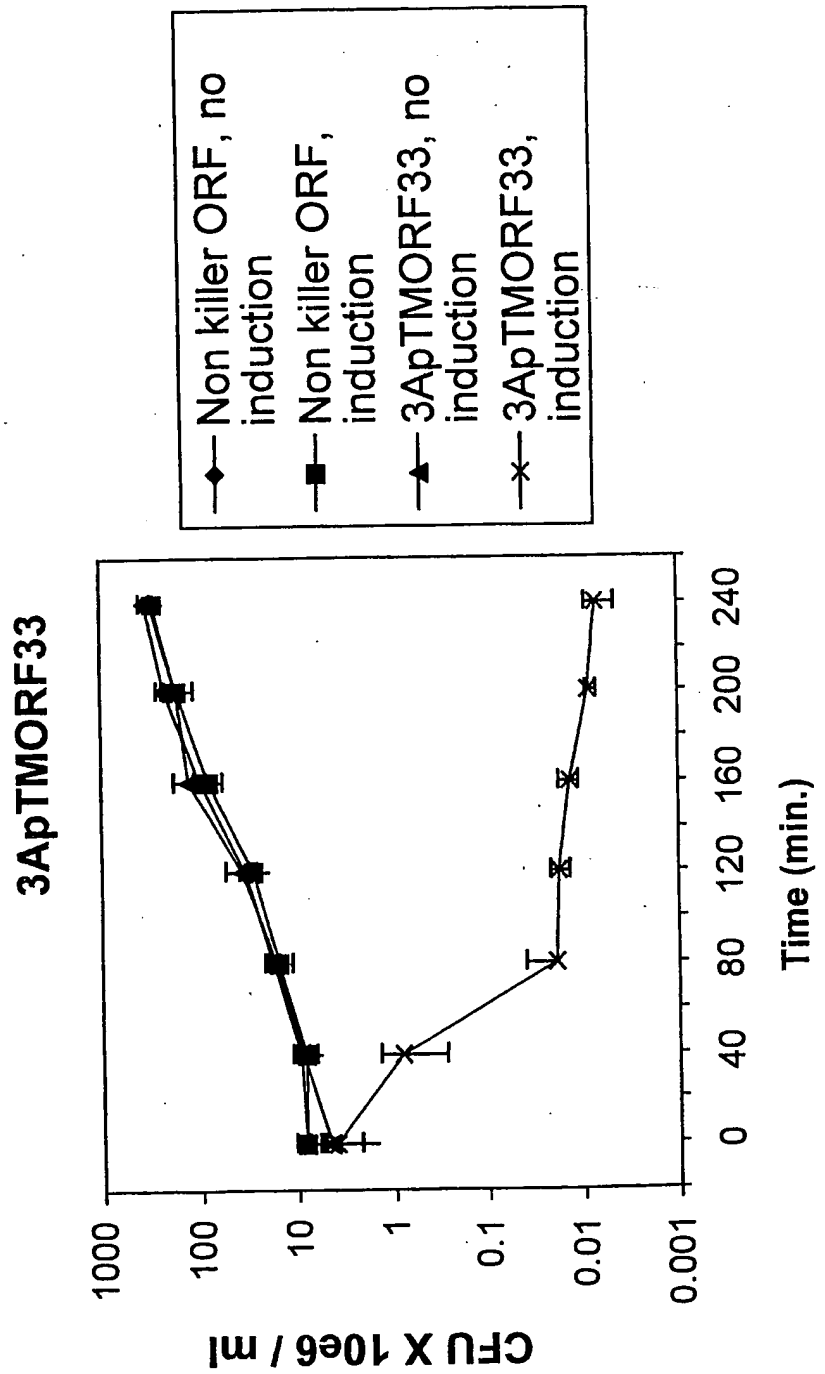

Non killer ORF, no induction
Non killer ORF, induction
3ApTMORF33, no induction
3ApTMORF33, induction
3ApTMORF33
CFU X 10e6 / ml
1000
100
10
1
0.1
0.01
0.001
0
40
80
120
160
200
240
Time (min.)

FIG. 3C

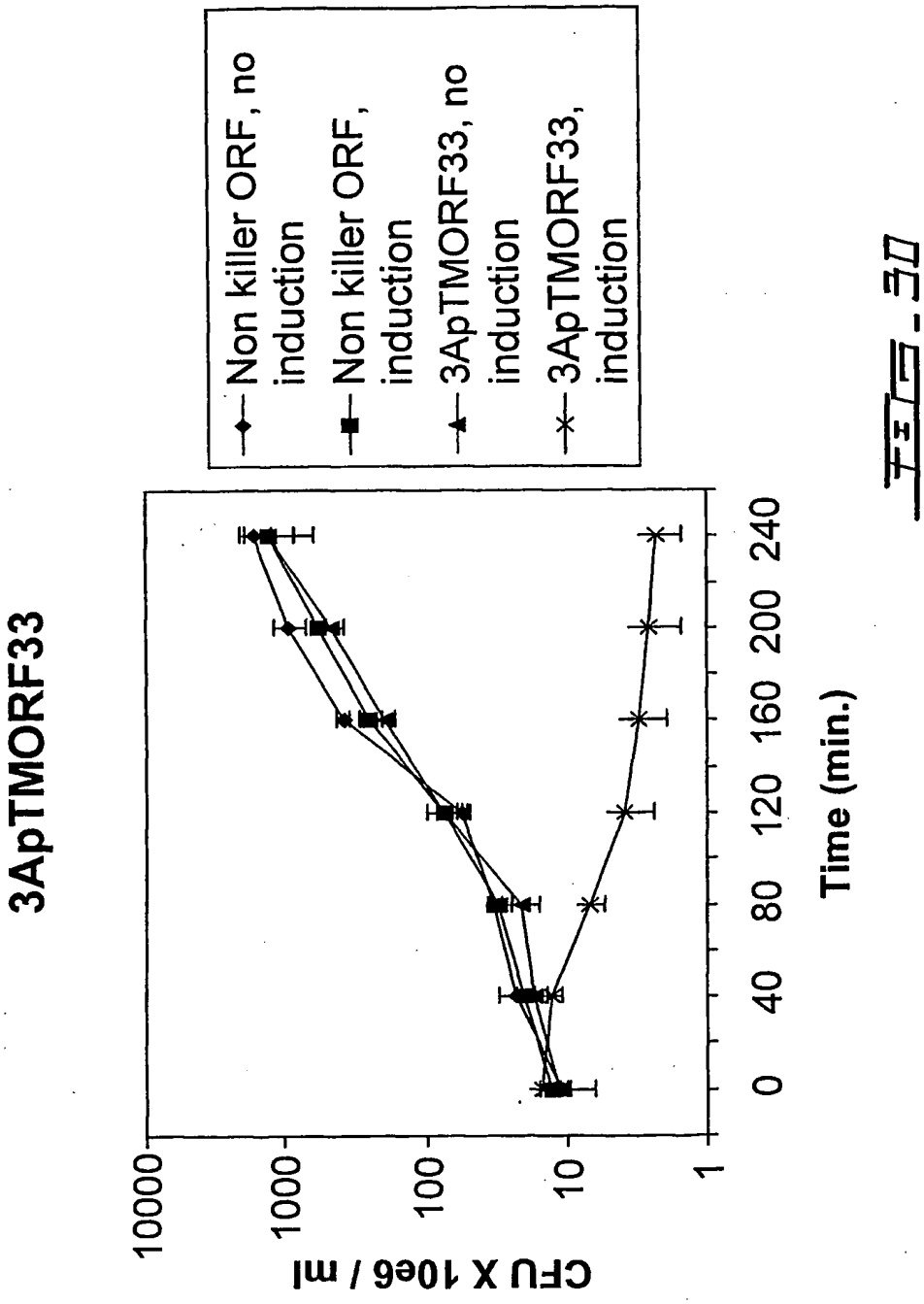

3ApTMORF33
CFU X 10e6 / ml
10000
1000
100
10
1
0
40
80
120
160
200
240
Time (min.)
Non killer ORF, no induction
Non killer ORF, induction
3ApTMORF33, no induction
3ApTMORF33, induction

FIG. 3D

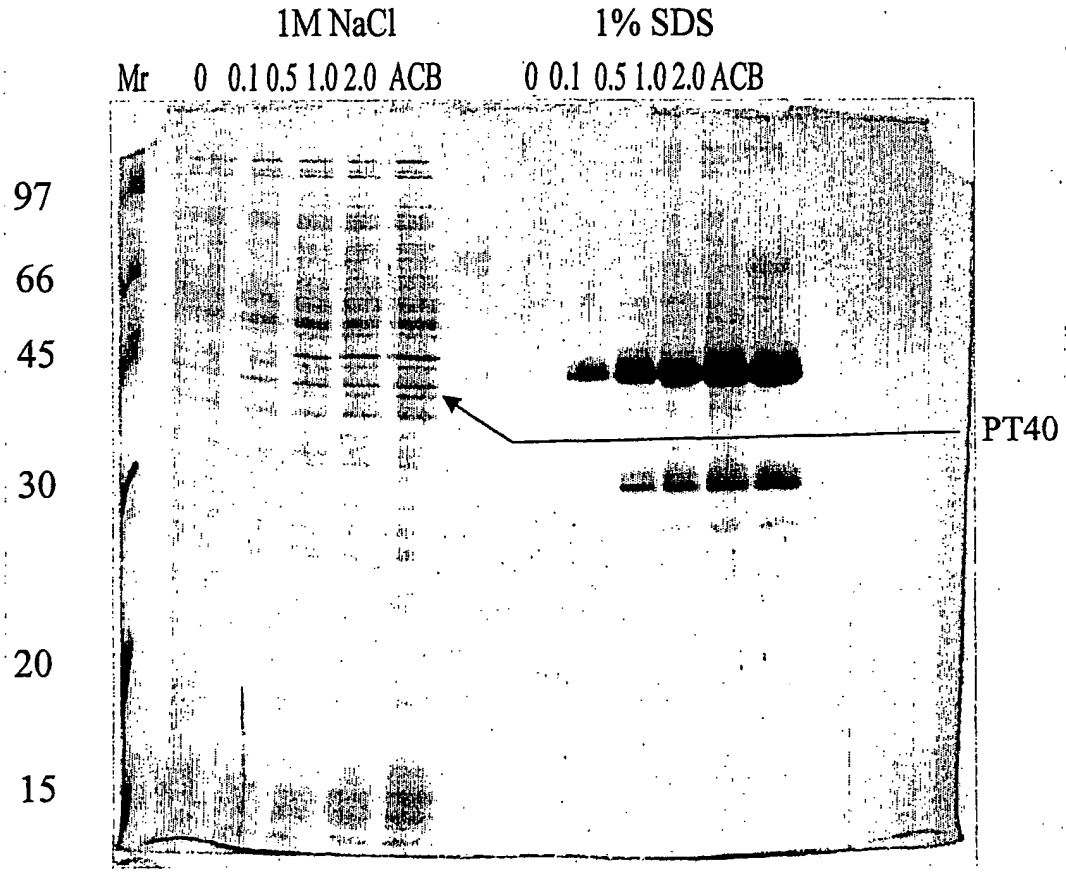

1M NaCl
1% SDS
Mr 0 0.1 0.5 1.0 2.0 ACB
0 0.1 0.5 1.0 2.0 ACB
97
66
45
30
20
15
PT40

FIG. 4

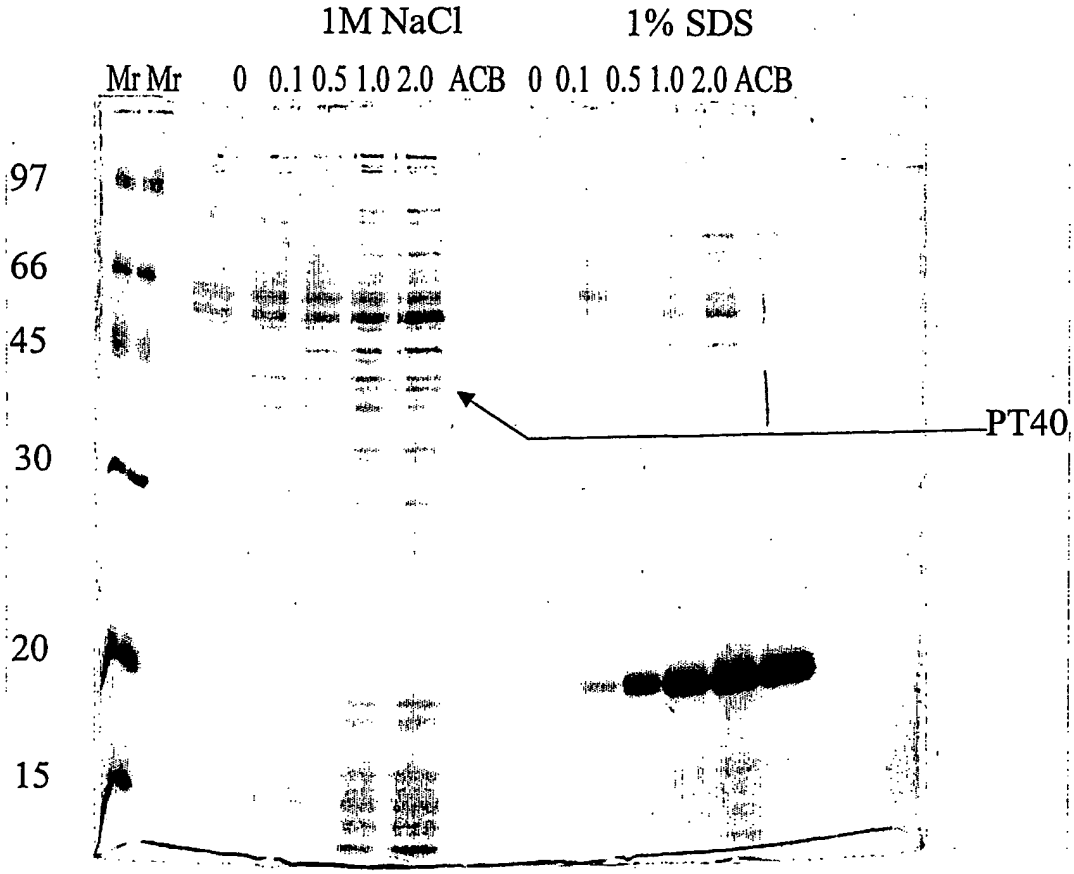

1M NaCl
1% SDS
Mr Mr 0 0.1 0.5 1.0 2.0 ACB 0 0.1 0.5 1.0 2.0 ACB
97
66
45
30
20
15
PT40

FIG. 5

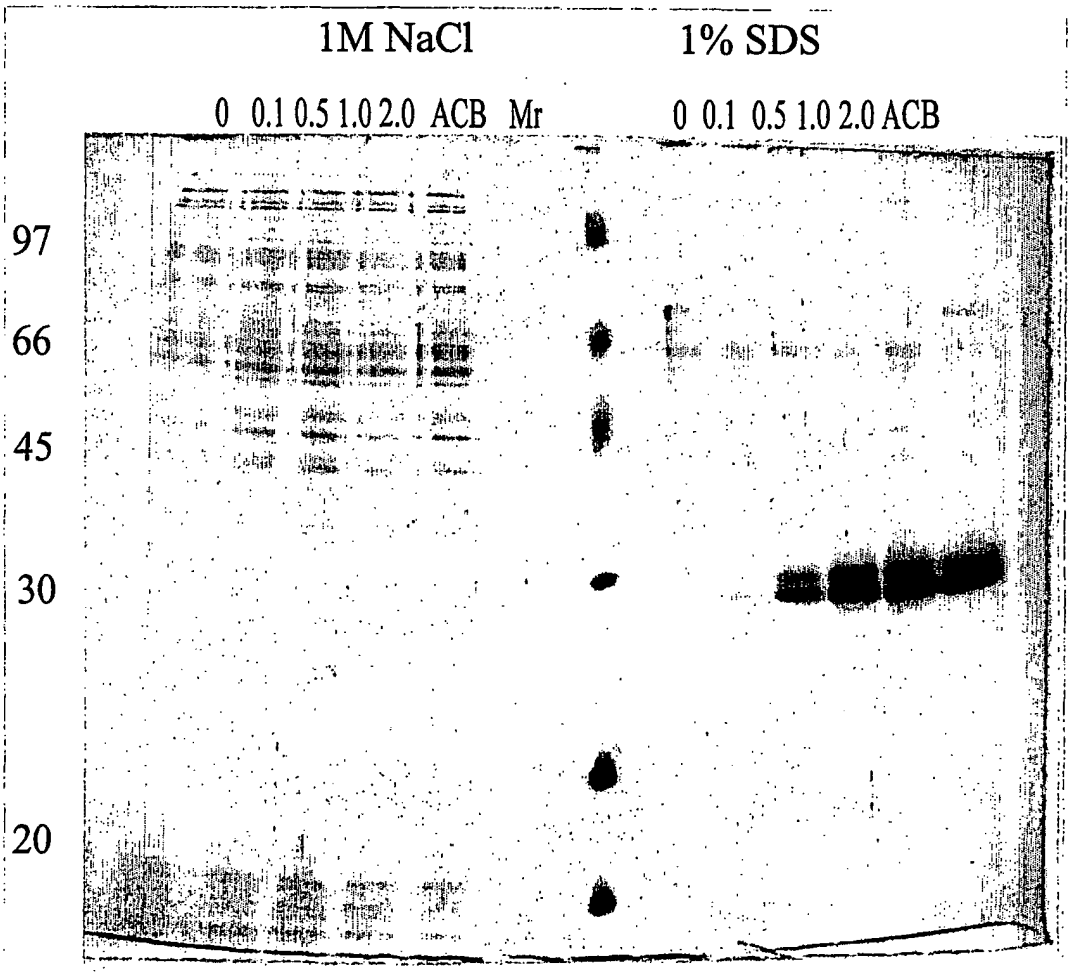


FIG. 6

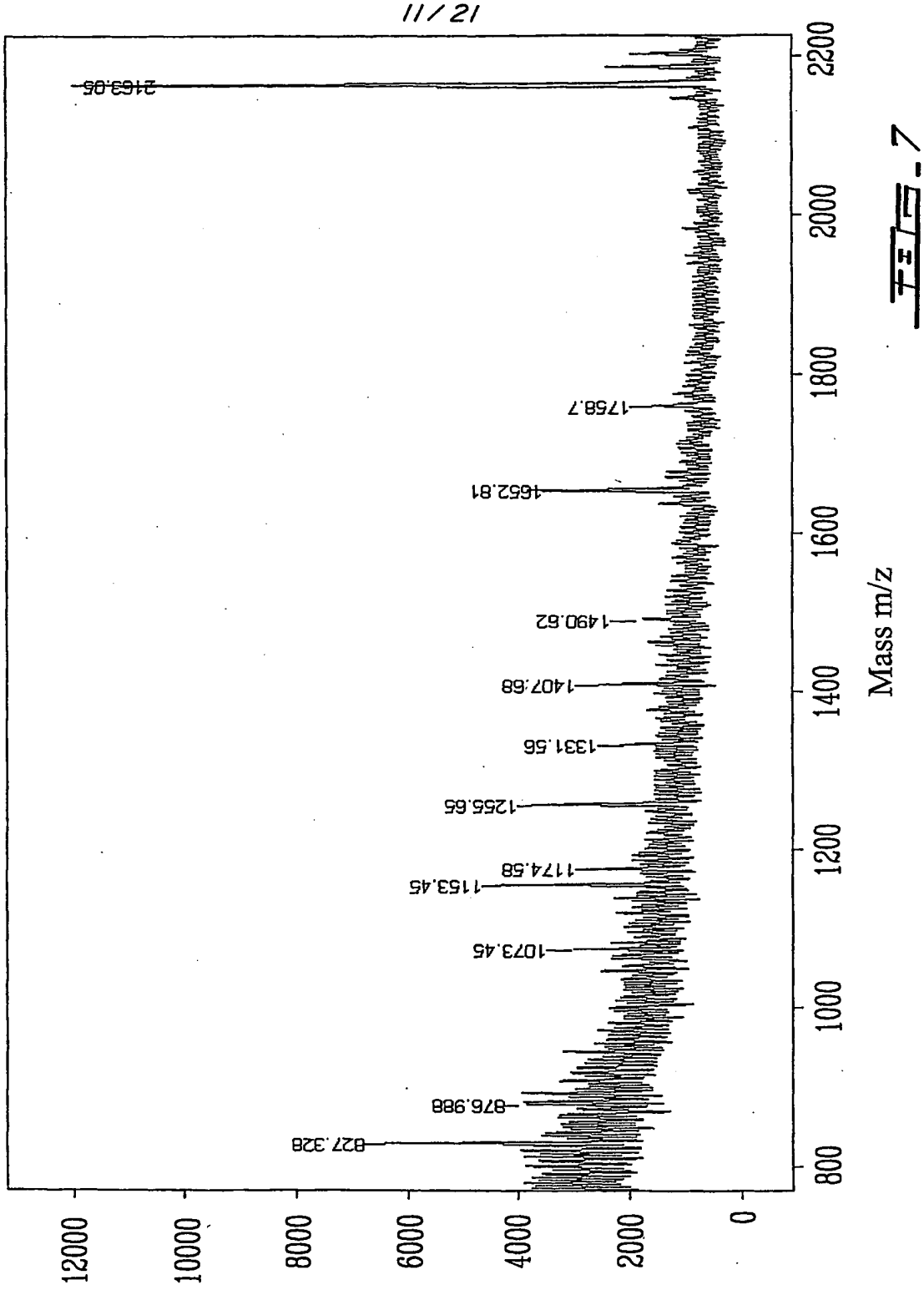

2163.05
1758.7
1652.81
1490.62
1407.68
1331.56
1255.65
1174.58
1153.45
1073.45
876.988
827.328
Mass m/z
800
1000
1200
1400
1600
1800
2000
2200
0
2000
4000
6000
8000
10000
12000

FIG. 7

Sequence analysis of STAAU_R004

Blast Analysis

Database: nr

582,290 sequences; 183,345,511 total letters

Sequences producing significant alignments:	Score (bits)	E Value
gi|6686325|sp|P71018|PLSX_BACSU FATTY ACID/PHOSPHOLIPID SYNTHES...	339	3e-92
gi|10175113|dbj|BAB06212.1| (AP001515) involved in fatty acid/p...	293	1e-78
gi|1502419|gb|AAC44305.1| (U59433) PlsX [Bacillus subtilis]	279	3e-74
gi|6685740|sp|O65984|PLSX_CLOTS FATTY ACID/PHOSPHOLIPID SYNTHES...	266	2e-70
gi|3006127|emb|CAA06178.1| (AJ004870) hypothetical protein [The...	259	2e-68
gi|6685741|sp|O67186|PLSX_AQUAE FATTY ACID/PHOSPHOLIPID SYNTHES...	221	9e-57
gi|6685761|sp|Q9WXZ6|PLSX_THEMA FATTY ACID/PHOSPHOLIPID SYNTHES...	211	8e-54
gi|7428758|pir||A64545 fatty acid/phospholipid synthesis protei...	210	1e-53
gi|6685765|sp|Q9ZMN1|PLSX_HELPJ FATTY ACID/PHOSPHOLIPID SYNTHES...	209	3e-53
gi|6685733|sp|O24993|PLSX_HELPY FATTY ACID/PHOSPHOLIPID SYNTHES...	208	5e-53
gi|2498794|sp|P73950|PLSX_SYNY3 FATTY ACID/PHOSPHOLIPID SYNTHES...	206	3e-52
gi|7441729|pir||C72412 fatty acid/phospholipid synthesis protei...	202	4e-51
gi|6967803|emb|CAB72796.1| (AL139074) putative fatty acid\phosp...	196	3e-49
gi|7379281|emb|CAB83835.1| (AL162753) putative fatty acid/phosp...	193	2e-48
gi|7227170|gb|AAF42243.1| (AE002540) fatty acid/phospholipid sy...	190	2e-47
gi|9656568|gb|AAF95172.1| (AE004276) fatty acid/phospholipid sy...	183	2e-45
gi|401571|sp|P30789|PLSX_RHOCA FATTY ACID/PHOSPHOLIPID SYNTHESI...	180	1e-44
gi|6685745|sp|O85138|PLSX_SALTY FATTY ACID/PHOSPHOLIPID SYNTHES...	178	7e-44
gi|6686317|sp|P27247|PLSX_ECOLI FATTY ACID/PHOSPHOLIPID SYNTHES...	178	9e-44
gi|9949068|gb|AAG06357.1|AE004723_1 (AE004723) fatty acid biosy...	176	3e-43
gi|2498793|sp|P75232|PLSX_MYCPN FATTY ACID/PHOSPHOLIPID SYNTHES...	175	6e-43
gi|2498791|sp|Q49427|PLSX_MYCGE FATTY ACID/PHOSPHOLIPID SYNTHES...	173	2e-42

FIG_8

Database: pdbaa

8891 sequences; 1,886,625 total letters

Sequences producing significant alignments:	Score (bits)	E Value
gi|4699699|pdb|1A40| Phosphate-Binding Protein With Ala 197 ...	26	5.2
gi|3892022|pdb|1A55|A Chain A, Phosphate-Binding Protein Mutant...	26	5.2
gi|2624641|pdb|1IXI| Phosphate-Binding Protein Mutant With A...	26	5.2
gi|2914192|pdb|1IXH| Phosphate-Binding Protein (Pbp) Complex...	26	5.2
gi|1633135|pdb|1QUI| Phosphate-Binding Protein Mutant With A...	26	5.2
gi|1633133|pdb|1QUK| Phosphate-Binding Protein Mutant With A...	26	5.2
gi|1633132|pdb|1QUL| Phosphate-Binding Protein Mutant With A...	26	5.2
gi|2914191|pdb|1IXG| Phosphate-Binding Protein Mutant With T...	26	5.2

Database: prodom

157,167 sequences; 18,560,502 total letters

Sequences producing significant alignments:	Score (bits)	E Value
pd|PD006974 p99.2 (12) PLSX(6) // PROTEIN PLSX SYNTHESIS FATTY...	266	2e-71
pd|PD017557 p99.2 (4) O34561(1) O65984(1) O67186(1) // PROTEIN...	52	5e-07
pd|PD136714 p99.2 (1) Q47878_EUBAC // GLYCINE REDUCTASE	44	1e-04
pd|PD123945 p99.2 (1) PLSX_RHOCA // FATTY ACID/PHOSPHOLIPID SYN...	37	0.018
pd|PD039564 p99.2 (2) PLSX(2) // FATTY ACID/PHOSPHOLIPID SYNTH...	35	0.093
pd|PD123943 p99.2 (1) O24993_HELPY // FATTY ACID/PHOSPHOLIPID S...	34	0.12
pd|PD131367 p99.2 (1) O61845_CAEEL // T04D1.4 PROTEIN	30	3.1
pd|PD131278 p99.2 (2) O17665(1) UN13(1) // UNC13 PROTEIN PHORB...	30	3.1
pd|PD181530 p99.2 (1) O84817_CHLTR // FA/PHOSPHOLIPID SYNTHESIS...	29	4.1
pd|PD014373 p99.2 (5) ACLY(3) P90731(1) Q13037(1) // LYASE ATP...	29	4.1
pd|PD141343 p99.2 (1) O60158_SCHPO // HYPOTHETICAL 49.4 KD PROT...	29	5.3

FIG. 8 (cont.)

Database: swissprot
88,166 sequences; 32,012,037 total letters

Sequences producing significant alignments:		Score (bits)	E Value
sp|P71018 PLSX_BACSU	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	339	5e-93
sp|O65984 PLSX_CLOTS	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	266	4e-71
sp|O67186 PLSX_AQUAE	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	221	2e-57
sp|Q9WXZ6 PLSX_THEMA	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	211	1e-54
sp|Q9ZMN1 PLSX_HELPJ	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	209	6e-54
sp|O24993 PLSX_HELPY	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	208	9e-54
sp|P73950 PLSX_SYNY3	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	206	5e-53
sp|P30789 PLSX_RHOCA	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	180	3e-45
sp|O85138 PLSX_SALTY	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	178	1e-44
sp|P27247 PLSX_ECOLI	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	178	2e-44
sp|P75232 PLSX_MYCPN	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	175	1e-43
sp|Q49427 PLSX_MYCGE	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	173	3e-43
sp|O84817 PLSX_CHLTR	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	145	9e-35
sp|Q9Z6U6 PLSX_CHLPN	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	142	1e-33
sp|Q46578 PLSX_DEIRA	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN...	138	2e-32

FIG. 8 (cont.)

Database: trembl

372,572 sequences; 116,471,635 total letters

Sequences producing significant alignments:		Score (bits)	E Value
tr|Q9KA01|Q9KA01	INVOLVED IN FATTY ACID/PHOSPHOLIPID SYNTHESIS.	293	9e-79
tr|Q9PIH0|Q9PIH0	PUTATIVE FATTY ACID\PHOSPHOLIPID SYNTHESIS PR...	196	2e-49
tr|Q9JW54|Q9JW54	PUTATIVE FATTY ACID/PHOSPHOLIPID BIOSYNTHESIS...	193	1e-48
tr|Q9JXR8|Q9JXR8	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN.	190	1e-47
tr|Q9KQH4|Q9KQH4	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLSX.	183	1e-45
tr|Q9RA35|Q9RA35	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLS...	168	4e-41
tr|Q9PQT7|Q9PQT7	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN.	153	1e-36
tr|Q9PLB1|Q9PLB1	FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN.	144	7e-34
tr|Q9JQ90|Q9JQ90	FA/PHOSPHOLIPID SYNTHESIS PROTEIN (FATTY ACID...	142	4e-33
tr|Q53843|Q53843	ORFA AND ORFF GENES, PARTIAL CDS, ORFB, ORFC,...	100	2e-20

FIG. 8 (cont.)

Database: staph_aureus
160 sequences; 5,603,166 total letters

Sequences producing significant alignments:	Score (bits)	E Value
st|8091 Staphylococcus aureus, TIGR database	647	0.0
st|Contig248 Staphylococcus aureus, Univ. of Oklahama database	647	0.0
st|8083 Staphylococcus aureus, TIGR database	28	1.9
st|Contig276 Staphylococcus aureus, Univ. of Oklahama database	28	1.9
st|8084 Staphylococcus aureus, TIGR database	28	2.5
st|Contig264 Staphylococcus aureus, Univ. of Oklahama database	28	2.5
st|8104 Staphylococcus aureus, TIGR database	28	2.5
st|8096 Staphylococcus aureus, TIGR database	27	4.3
st|Contig280 Staphylococcus aureus, Univ. of Oklahama database	27	4.3
st|8085 Staphylococcus aureus, TIGR database	26	9.6

FIG_8 (cont.)

17/21

Optimal global alignment between STAAU_R004 and gene PlsX from *Bacillus subtilis*

Sequence 1

>pt|125098 STAAU_R004 STAAU_R004_NT|1-987|
(328 letters)

Sequence 2

>sp|P71018 PLSX_BACSU FATTY ACID/PHOSPHOLIPID SYNTHESIS PROTEIN PLSX.
(333 letters)

Substitution matrix: blosum62
Gap penalty: - (11 + 1 * (gap length))

Score: 908
Identical: 179/334 (53%), **Similar:** 240/334 (71%), **Gap:** 7/334 (2%)

```
seq1     1 MVKLAIDMMGGDNAPDIVLEAVQKAVEDFKDLEIILFGD----EKKYNLNHERIEFRHCS   56
           | ::|:| ||||:||  |:: | | :| | || | | ||    |       :||   |
seq2     1 M-RIAVDAMGGDHAPKAVIDGVIKGIEAFDDLHITLVGDKTTIESHLTTTSDRITVLHAD   59

seq1    57 EKIEMEDEPVRAIKRKKDSSMVKMAEAVKSGEADGCVSAGNTGALMSAGLFIVGRIKGVA   116
           | ||  ||||||::|||:|||| ||: |    || |:||||||||:||||||||||:
seq2    60 EVIEPTDEPVRAVRRKKNSSMVLMAQEVAENRADACISAGNTGALMTAGLFIVGRIKGID   119
```

FIG. 9A

```
seq1   117 RPALVVTLPTIDGKGFVFLDVGANADAKPEHLLQYAQLGDIYAQKIRGIDNPKISLLNIG   176
           ||||  ||||: | ||: |||||| |||||||:||| :| :|:|::||: :|:: |||:|
seq2   120 RPALAPTLPTVSGDGFLLLDVGANVDAKPEHLVQYAIMGSVYSQQVRGVTSPRVGLLNVG   179

seq1   177 TEPAKGNSLTKKSYELLNHDHSLNFVGNIEAKTLMDGDTDVVVTDGYTGNMVLKNLEGTA   236
           ||  ||| |||:::::|    ::||:||:||: |:|   |||||||:|||: || |||:|
seq2   180 TEDKKGNELTKQTFQILKETANINFIGNVEARDLLDDVADVVVTDGFTGNVTLKTLEGSA   239

seq1   237 KSIGKMLKDTIMSSTKNKLAGAILKKDLAEFAKKMDYSEYGGSVLLGLEGTVVKAHGSSN   296
            || ||::| : |:  :||| |:||  | |   ||:|| |||: | ||:  |:||||||:
seq2   240 LSIFKMMRDVMTSTLTSKLAAAVLKPKLKEMKMKMEYSNYGGASLFGLKAPVIKAHGSSD   299

seq1   297 AKAFYSAIRQAKIAGEQNIVQTMKETVGE--SNE   328
           : | : |||||:    ||:   ::| | |  ::|
seq2   300 SNAVFHAIRQAREMVSQNVAALIQEEVKEEKTDE   333
```

FIG. 9A (cont.)

STAAU_R004 stop codon-to-stop codon region

```
50574  tagttttcaaaggaaattttaaaatgttttatgataagcgaggataaaattatggttaaattagcaatt
1      * F  S  K  E  I  L  K  C  F  M  I  S  E  D  K  I  M  V  K  L  A  I
50643  gatatgatgggtggcgacaatgcgcctgatatcgtattagaagccgtacaaaaggctgttgaagacttt
24     D  M  M  G  G  D  N  A  P  D  I  V  L  E  A  V  Q  K  A  V  E  D  F
50712  aaagatctagaaattatacttttcggtgacgaaaaaaagtataatctgaaccatgaacgaatcgaattt
47     K  D  L  E  I  I  L  F  G  D  E  K  K  Y  N  L  N  H  E  R  I  E  F
50781  agacattgttctgaaaagattgaaatggaagatgagcctgttagagcgattaaacgtaaaaaagatagc
70     R  H  C  S  E  K  I  E  M  E  D  E  P  V  R  A  I  K  R  K  K  D  S
50850  tcaatggtaaaaatggctgaagctgtgaaatctggtgaagcagatggatgtgtgtcagcaggtaatact
93     S  M  V  K  M  A  E  A  V  K  S  G  E  A  D  G  C  V  S  A  G  N  T
50919  ggtgctttaatgtcagctggtttattcattgttggacgtattaaaggtgtagctagaccggctttagta
116    G  A  L  M  S  A  G  L  F  I  V  G  R  I  K  G  V  A  R  P  A  L  V
50988  gtaacattgccaacgattgatggaaaaggttttgtctttttagacgttggtgcaaatgctgatgctaaa
139    V  T  L  P  T  I  D  G  K  G  F  V  F  L  D  V  G  A  N  A  D  A  K
51057  cctgaacacttattacagtatgcgcaactaggggatatttatgctcaaaaaattagaggtattgataat
162    P  E  H  L  L  Q  Y  A  Q  L  G  D  I  Y  A  Q  K  I  R  G  I  D  N
51126  ccgaaaatctcattattaaatataggaaccgagccagctaaaggtaatagtttaacgaaaaaatcatat
185    P  K  I  S  L  L  N  I  G  T  E  P  A  K  G  N  S  L  T  K  K  S  Y
51195  gagttattaaatcatgatcattcattgaattttgttgggaatattgaagcgaagacattaatggatggc
208    E  L  L  N  H  D  H  S  L  N  F  V  G  N  I  E  A  K  T  L  M  D  G
51264  gatacagatgttgtagttaccgatggctatactgggaacatggtccttaaaaatttagaaggtactgca
231    D  T  D  V  V  V  T  D  G  Y  T  G  N  M  V  L  K  N  L  E  G  T  A
51333  aaatcaatcggtaaaatgttaaaagatacgattatgagtagtactaaaaataaattagcaggtgcaata
254    K  S  I  G  K  M  L  K  D  T  I  M  S  S  T  K  N  K  L  A  G  A  I
51402  ttgaagaaagatttagctgaattcgctaaaaagatggattactcagaatacggtggttccgtattatta
277    L  K  K  D  L  A  E  F  A  K  K  M  D  Y  S  E  Y  G  G  S  V  L  L
51471  ggattggaaggtactgtagttaaagcacacggtagttcaaatgctaaagcttttttattctgcaattaga
300    G  L  E  G  T  V  V  K  A  H  G  S  S  N  A  K  A  F  Y  S  A  I  R
51540  caagcgaaaatcgcaggagaacaaaatattgtacaaacaatgaaagagactgtaggtgaatcaaatgag
323    Q  A  K  I  A  G  E  Q  N  I  V  Q  T  M  K  E  T  V  G  E  S  N  E
51609  taa 51611
346    *
```

FIG_9B

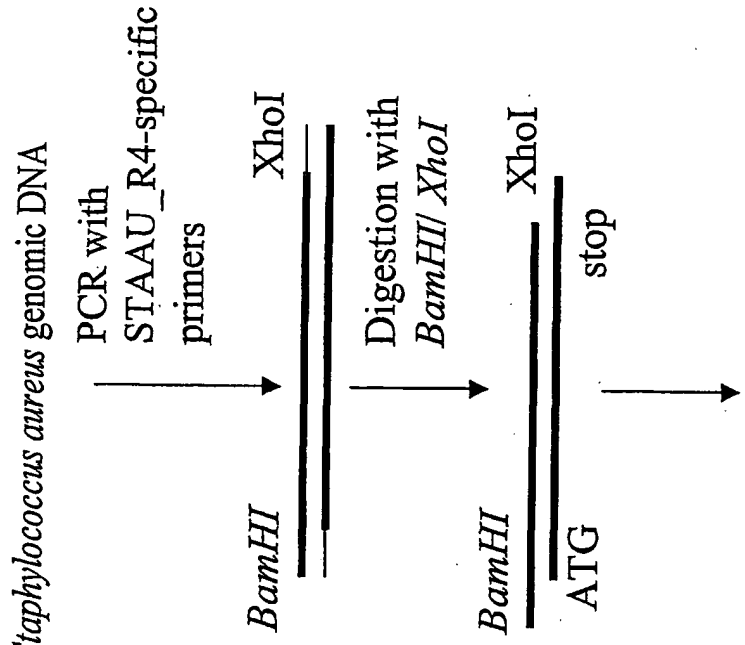


FIG - 10A

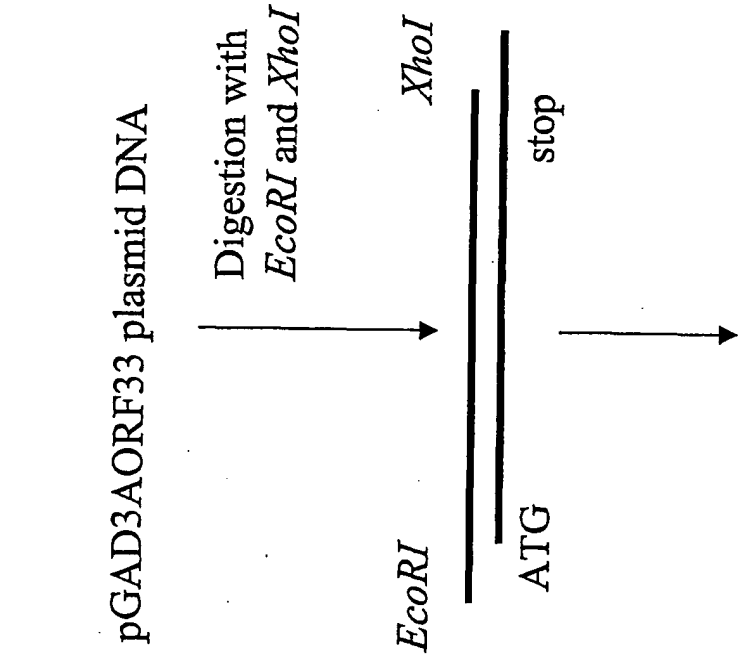


FIG - 10B

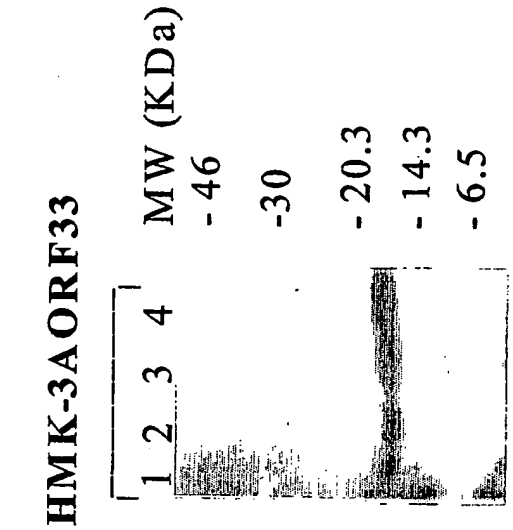


FIG. 10D

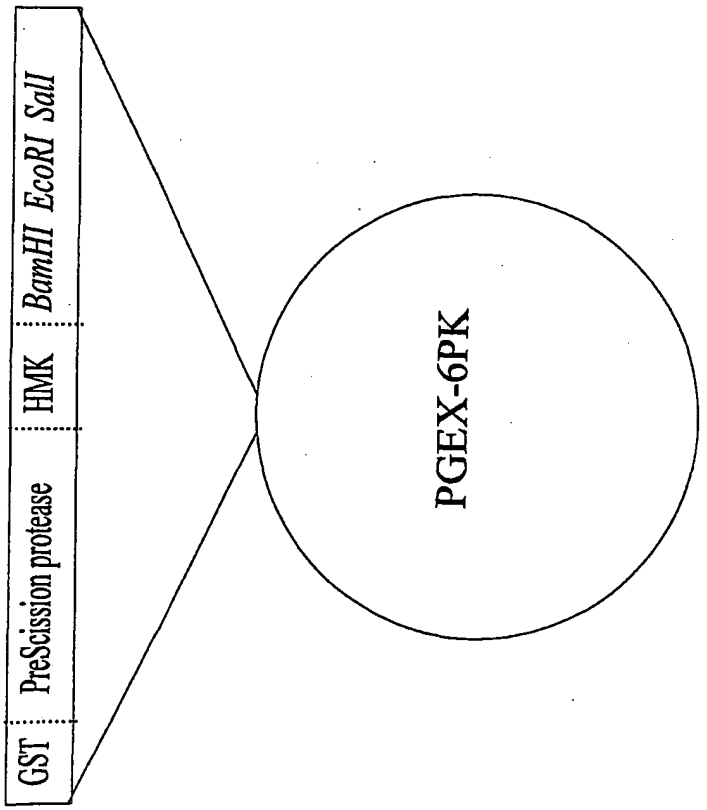


FIG. 10C